(54) Title: OLIGONUCLEOTIDE IDENTIFIERS

(57) Abstract: Methods, assays, and components are described in which biological samples can be rapidly and sensitively analyzed for the presence of species associated with neurodegenerative disease. Techniques and components are provided for diagnosis of disease, as well as for screening of candidate drugs for treatment of neurodegenerative disease. The techniques are simple, extremely sensitive, and utilize readily-available components. Binding species, capable of binding a neurodegenerative disease aggregate-forming or aggregate-forming species, are fastened to surfaces of electrodes and surfaces of particles, or provided free in solution, to bind aggregate-forming species and/or be involved in aggregation.

# OLIGONUCLEOTIDE IDENTIFIERS

Related Applications

This application claims priority to U.S. Provisional Patent Application Serial

5 Numbers 60/248,863, filed November 15, 2000, 60/252,650, filed November 22, 2000,

Great Britain Patent Application Serial Number 0101054.5, filed January 15, 2001,

U.S. Provisional Patent Application Serial Numbers 60/276,995, filed March 19, 2001,

60/302,231, filed June 29, 2001, 60/326,937, filed October 3, 2001, and 60/327,089,

filed October 3, 2001.

10

Field of the Invention

This invention relates generally to methods, assays, and components for the

rapid, high-throughput, specific and sensitive detection and analysis of biomolecular

and chemical interactions, and more particularly to identifiers for identification of

15 participants in these and other assays.

Background of the Invention

International Patent Application Serial No. PCT/US00/01504, published July

27, 2000 as WO 00/43783 describes a variety of assays involving colloids.

International Patent Publication No. PCT/US00/01997, filed January 25, 2000

20 and U.S. Patent Application Serial No. 09/631,818, filed August 3, 2000 describe

methods, assays, and components for analyzing species associated with disease and for

screening of candidate drugs for treatment of disease. Assays involving colloid/colloid

interaction are described in detail.

*In vitro* techniques that currently exist for studying protein-protein interactions

25 include co-immunoprecipitation, co-fractionation by chromatography, cross-linking,

sandwich assays and surface plasmon resonance. A disadvantage of these techniques is

that putative binding partners must be sequentially tested which greatly limits the

number of potential interacting proteins that can be tested. There are two reasons why

experiments have to be performed sequentially. The first is a signaling problem. In a

30 typical binding assay, a single type of signal is produced when binding occurs. Putative

binding partners must therefore be kept isolated, then tested sequentially for pair-wise

interactions. The second reason for sequential experiments is a bookkeeping problem.

Since it is very difficult to identify proteins, especially when at low concentrations, it is

- 2 -

necessary to keep track of isolated and purified species, then test for binding in pair-wise fashion.

*In vivo*, cell-based binding assays such as the yeast two hybrid system and yeast mating system provide a major advantage over existing *in vitro* methods in that once a positive protein-protein interaction has been detected, the host cell, which provides the signal, contains ample DNA that codes for the protein(s) under study. As those skilled in the art appreciate, it is far easier to sequence DNA that proteins or peptides. Additionally, DNA at low concentration can be enzymatically amplified prior to sequencing whereas proteins cannot. This eliminates the need for tracking individual aliquots of purified proteins and facilitates high throughput screening to detect protein-protein interactions.

There are, however disadvantages of *in vivo* protein detection systems. *In vivo* assays suffer from false positives and negatives because of the inherent redundancies of the biological processes upon which the assays are based. For example, the yeast two-hybrid system is based on a mechanism of transcriptional activation. Another disadvantage of the system is that it can only detect interactions between cell-derived species. Therefore, interactions between proteins and chemical species, such as drug candidates or chemical recognition elements cannot be detected using this method.

While a wide variety of biological and chemical assay techniques are known, assays with enhanced multiplexing capability that do not sacrifice accuracy in detection would be advantageous. Therefore, an *in vitro* binding assay in which genetic material that codes for expressed proteins or chemical species is available, and can be correlated to a specific species after the binding assay, would provide for high throughput and a major advantage over existing systems.

Summary of the Invention

The present invention provides a series of methods, components, kits, etc. for use in chemical and biological analysis. Specifically, the invention provides techniques for studying binding interactions between chemical or biological species, such as binding interactions between proteins. The invention allows for high-throughput, multiplexed screening of interactions between species. That is, large numbers of interactions can be screened simultaneously, as opposed to those prior techniques in which binding partner candidates were screened sequentially. It is a significant

advantage of the present invention that high-throughput, multiplexed screening can be conducted *in vitro*.

The invention provides techniques for determination of where binding interactions have occurred among many possibilities of binding interactions, and rapid
5    selection and identification of species that have participated in binding interactions.

Another method involves allowing a chemical or biological species, immobilized relative to a surface, to participate in a chemical or biological interaction. The identification of the chemical or biological species that participated in the interaction is then determined by identifying an oligonucleotide identifier associated
10    with the surface, optionally by identifying a unique combination of two oligonucleotide identifiers associated with each of the interacting partners. Identifying the combined identifiers uniquely identifies the interacting pair. In one embodiment, an identifier can be an oligonucleotide that codes for a protein that it identifies in the assay.

Another method involves expressing a protein with an oligonucleotide, and
15    immobilizing the protein and the oligonucleotide relative to each other.

Another method of the invention includes expressing a protein from a nucleic acid and immobilizing the protein and the oligonucleotide relative to each other.

Another aspect of the invention involves articles. One article of the invention has a surface with a chemical or biological species able to participate in a chemical or
20    biological interaction, fastened to or adapted to be fastened to the surface. An oligonucleotide identifier is also fastened to or adapted to be fastened to the surface. As in the above method, in this aspect, according to one embodiment, the identifier can be an oligonucleotide that codes for a protein that it identifies in the assay.

The invention also provides kits for biological or chemical analysis. One kit is
25    defined as the article described in the paragraph above. Another kit includes at least one additional article having a surface, a second chemical or biological species fastened to or adapted to be fastened to the surface, and a second oligonucleotide identifier fastened to or adapted to be fastened to the surface. The kit can contain a set of oligonucleotides, or derivatives of oligonucleotides, which completely represent all
30    possible oligonucleotide identifier sequence combinations that uniquely identify any two interacting partners. The kit can contain a set of nucleic acid identifiers to interactions that involve more than two interacting species. In another embodiment a

- 4 -

kit of the invention includes a surface, a protein immobilized relative to the surface or adapted to be immobilized relative to the surface, and an oligonucleotide identifier that codes for the protein, immobilized relative to the surface or adapted to be immobilized relative to the surface.

In another embodiment a kit of the invention includes a polymer or dendrimer rather than a surface.

In another embodiment a kit of the invention includes a protein and an oligonucleotide identifier that codes for the protein, immobilized relative to each other or adapted to be immobilized relative to each other, and an entity carrying immobilized thereto a binding partner of the protein.

Another aspect of the invention involves compositions. One composition comprises a protein and an oligonucleotide identifier that codes for the protein, immobilized relative to each other or adapted to be immobilized relative to each other.

Another composition of the invention comprises a chemical or biological species, able to participate in a chemical or biological interaction, a linker species that is not a ribosome, and an oligonucleotide identifier, wherein each of the chemical or biological species and the oligonucleotide identifier is fastened to or adapted to be fastened to the linker species. The linker species can be a surface of nanoparticle, chip, polymer, dendrimer, RNA binding protein, DNA binding protein, etc.

It is not intended that the present invention be limited by the nature of the solid support. In one embodiment, the solid support is a colloid (e.g. gold colloid). It is also not intended that the present invention be limited by the nature of attachment of the ligand to the solid support. In one embodiment, said ligand is covalently attached (directly or through another ligand or binding moiety) to the solid support. In another embodiment, the ligand is attached non-covalently or by electrostatic or ionic interaction.

In some embodiments signaling entities are useful. In such embodiments, the invention contemplates a variety of signaling entities described below, including but not limited to optically active entities such as fluorescent molecules and enzymes capable of acting on color-producing substrates. Preferred signaling entities include electroactive molecules, that is, molecules having an oxidation/reduction potential that

can be determined electronically or electrochemically proximate a working electrode of an appropriate, conventional electrical arrangement, as signaling elements.

It is not intended that the present invention be limited by the nature of the chemical or biochemical agent. A wide variety of agents and binding partners of those agents such as protein/protein, protein/peptide, antibody/antigen, antibody/hapten, enzyme/substrate, enzyme/inhibitor, enzyme/cofactor, binding protein/substrate, carrier protein/substrate, lectin/carbohydrate, receptor/hormone, receptor/effector, complementary strands of nucleic acid, protein/nucleic acid repressor/inducer, ligand/cell surface receptor, virus/ligand, etc., can be used for binding interactions of the inventions. In one embodiment, the agent is a ligand, specifically a peptide. In a preferred embodiment, the peptide is derivatized with a moiety (such as a histidine tag) that can bind to a metal chelate. In this embodiment, it is convenient that the solid support comprise a metal chelate and said peptide is attached to said solid support via binding of said moiety to said metal chelate.

In some embodiments, cell-derived molecules, including both cell-surface receptors and intracellular signaling proteins, exist on or are attached to solid supports that can either be surfaces or particle-like in nature. Binding partners of these cell-derived proteins, which can include both known and unknown ligands as well as putative drug candidates, are attached to surfaces and/or particle-like structures, and are allowed to interact with the cell-derived proteins in a manner such that binding between the two binding partners occurs. One of the binding partners or its attached support can additionally be derivatized with a detectable substance. Interacting complexes are identified using characteristics of the associated complex that differentiate it from the unassociated binding partners. The presence of, or a change in, a detectable moiety, that is either co-immobilized with one of the binding partners on a common solid support or directly attached to one of the binding partners, is detected. Molecules that disrupt a relevant interaction can be identified by detecting a loss of this signal. Interacting partners are brought to a sensing apparatus by confining one of the binding partners to the sensing area and allowing it to recruit the other binding partner, or by manipulating characteristics of the associated complex that differentiate it from the unassociated binding partners, or by attaching a recruitable element to one of the binding partners or its associated solid support.

- 6 -

One embodiment of the invention involves recruiting an electronic signaling entity to an electrode using a magnetic material. This embodiment can find use in many assays and other techniques of the invention. In the method, typically, a signaling entity is provided with the ability to become immobilized relative to the magnetic material (which can be a magnetic bead). The magnetic material and signaling entity can become immobilized relative to each other via a variety of chemical and/or physical linkages described herein. For example a first species may be immobilized relative to or fastened to a magnetic material and a second species may be immobilized relative to or fastened to the signaling entity or the first and second species can bind to each other. The first and second species can be essentially any species described herein or known in the art for binding, and in one embodiment are proteins. In a preferred embodiment, the proteins are not antibodies but are, for example, a ligand and a cognate receptor, etc. The signaling entity can be fastened to an intermediate entity, such as a colloid particle, to which one of the proteins that acts as a binding partner also is fastened. A signaling entity also can be recruited to an electrode without use of a magnetic material. In this arrangement the signaling entity can be immobilized with respect to a binding partner of a species with respect to the electrode, and the binding partners can be allowed to bind to each other. Essentially any binding partner interaction as described herein or known in the art can facilitate this technique.

It is to be understood that aspects of the invention involving an oligonucleotide identifier can be used in connection with any aspects described herein, and that the oligonucleotide identifier has application in essentially any chemical or biological binding study.

Other advantages, novel features, and objects of the invention will become apparent from the following detailed description of the invention when considered in conjunction with the accompanying drawings, which are schematic and which are not intended to be drawn to scale. In the figures, each identical or nearly identical component that is illustrated in various figures is represented by a single numeral. For purposes of clarity, not every component is labeled in every figure, nor is every component of each embodiment of the invention shown where illustration is not necessary to allow those of ordinary skill in the art to understand the invention.

- 7 -

Brief Description of the Drawings

Fig. 1 illustrates schematically an embodiment of a colloid particle 140 adapted to bind essentially any chemical or biological species and also to bind an oligonucleotide identifier.

5       Fig. 2 illustrates schematically a chip including a plurality of spatially-addressable regions, each region having a chemical or biological species (putative binding species) and an oligonucleotide identifier.

Fig. 3 illustrates, schematically, another embodiment showing a chip to which one or more chemical or biological species are fastened.

10      Fig. 4 illustrates an oligonucleotide identifier of the invention adapted to be fastened to a surface, specifically via a self-assembled monolayer- forming species.

Fig. 5 illustrates identification of the polyamino acid tag of Figures 4-8, following separation from the surface of the colloid particle to which it had been fastened.

15      Fig. 6 illustrates a surface of a colloid particle to which is fastened an oligonucleotide identifier (Fig. 4) and a biological binding partner.

Fig. 7 illustrates biological binding between first and second biological binding partners attached to first and second colloid particles, respectively.

Fig. 8 illustrates separation of the oligonucleotide identifier of Fig. 6 from the

20      surface of the colloid particle to which it had been fastened.

Fig. 9 illustrates an oligonucleotide identifier and a biological binding partner, each fastened to a surface of a colloid particle.

Fig. 10 illustrates two colloid particles, each carrying a biological species that biologically binds to the species of the other colloid particle, and each carrying a

25      oligonucleotide identifier.

Fig. 11 illustrates binding of an interaction hybridization identifier to the combination of the oligonucleotide identifiers bound, respectively, to the colloid particles of Fig. 10.

Fig. 12 illustrates de-activating any non-hybridized oligonucleotide.

30      Fig. 13 illustrates the result of the step of Fig. 12.

Fig. 14 illustrates denaturization of the interaction hybridization identifier of Figures 11-13;

- 8 -

Fig. 15 illustrates identification of chimeric oligo solution and thereby identification of the oligonucleotide identifiers of Figures 10-13.

Fig. 16 shows ACV demonstration of enhanced electronic communication across a self-assembled monolayer, and redox signaling of protein immobilization to a cell surface, against a control.

Fig. 17 shows ACV analysis of protein/protein interaction as measured by binding of a colloid to a magnetic bead.

Fig. 18 illustrates how two binding partners can be detected through magnetic recruitment.

Fig. 19 illustrates a multiplexing apparatus for applying and releasing a magnetic force at multiple locations on a continuous surface.

Detailed Description of the Invention

International patent application serial number PCT/US00/01997, filed 01/25/00 by Bamdad et al., entitled "Rapid and Sensitive Detection of Aberrant Protein Aggregation in Neurodegenerative Diseases" (published as WO 00/43791 on 07/27/00), International patent application serial number PCT/US00/01504, filed 01/21/00 by Bamdad, et al, entitled "Interaction of Colloid-Immobilized Species with Species on Non-Colloidal Structures" (published as WO 00/34783 on 07/27/00), commonly-owned, copending U.S. patent application serial no. 09/602,778, filed 06/23/00 by Bamdad et al., entitled "Interaction of Colloid-Immobilized Species with Species on Non-Colloidal Structures"; and commonly-owned, copending U.S. patent application serial no. 09/631,818, filed 08/03/00 by Bamdad et al., entitled "Rapid and Sensitive Detection of Protein Aggregation" all are incorporated herein by reference.

"Small molecule", as used herein, means a molecule less than 5 kiloDalton, more typically less than 1 kiloDalton. As used herein, "small molecule" excludes proteins.

The term "candidate drug" as used herein, refers to any medicinal substance used in humans, animals, or plants. Encompassed within this definition are compound analogs, naturally occurring, synthetic and recombinant pharmaceuticals, hormones,

antimicrobials, neurotransmitters, etc. This includes any substance or precursor (whether naturally occurring, synthetic or recombinant) which is to be evaluated for use as a drug for treatment of neurodegenerative disease, or other disease characterized by aberrant aggregation, or prevention thereof. Evaluation typically takes place through activity in an assay, such as the screening assays of the present invention.

A variety of types of particles can be used in the invention. For example, "fluid suspendable particle" means a particle that can be made to stay in suspension in a fluid in which it is used for purposes of the invention (typically an aqueous solution) by itself, or can be maintained in solution by application of a magnetic field, an electromagnetic field, agitation such as stirring, shaking, vibrating, sonicating, centrifuging, vortexing, or the like. Examples include colloid particles, nanocrystals, and the like. A "nanoparticle" is a particle that can be fluid-suspendable, having a maximum cross-sectional dimension of no more than 500 nanometers, preferably no more than 250 nanometers. A "magnetically suspendable" particle is one that can be maintained in suspension in a fluid via application of a magnetic field. An electromagnetically-suspendable particle is one that can be maintained in suspension in a fluid by application of an electromagnetic field (e.g., a particle carrying a charge, or a particle modified to carry a charge). A "self-suspendable particle" is a particle that is of low enough size and/or mass that it will remain in suspension in a fluid in which it is used (typically an aqueous solution), without assistance of for example a magnetic field, for at least 1 hour. Other self-suspendable particles will remain in suspension, without assistance, for 5 hours, 1 day, 1 week, or even 1 month, in accordance with the invention.

"Proteins" and "peptides" are well-known terms in the art, and are not precisely defined in the art in terms of the number of amino acids that each includes. As used herein, these terms are given their ordinary meaning in the art. Generally, peptides are amino acid sequences of less than about 100 amino acids in length, but can include sequences of up to 300 amino acids. Proteins generally are considered to be molecules of at least 100 amino acids.

As used herein, a "metal binding tag" refers to a group of molecules that can become fastened to a metal that is coordinated by a chelate. Suitable groups of such molecules include amino acid sequences, typically from about 2 to about 10 amino acid

- 10 -

residues. These include, but are not limited to, histidines and cysteines ("polyamino acid tags"). Such binding tags, when they include histidine, can be referred to as a "poly-histidine tract" or "histidine tag" or "HIS-tag", and can be present at either the amino- or carboxy-terminus, or at any exposed region, of a peptide or protein or nucleic

5    acid. A poly-histidine tract of six to ten residues is preferred for use in the invention. The poly-histidine tract is also defined functionally as being a number of consecutive histidine residues added to a protein of interest which allows the affinity purification of the resulting protein on a metal chelate column, or the identification of a protein terminus through the interaction with another molecule (e.g. an antibody reactive with

10   the HIS-tag).

"Affinity tag" is given its ordinary meaning in the art. Affinity tags include, for example, metal binding tags, GST (in GST/glutathione binding clip), and streptavidin (in biotin/streptavidin binding). At various locations herein specific affinity tags are described in connection with binding interactions. It is to be understood that the

15   invention involves, in any embodiment employing an affinity tag, a series of individual embodiments each involving selection of any of the affinity tags described herein.

As used herein, "chelate coordinating a metal" or metal coordinated by a chelate, refers to a metal coordinated by a chelating agent that does not fill all available coordination sites on the metal, leaving some coordination sites available for binding

20   via a metal binding tag. U.S. Patent No. 5,620,850 of Bamdad, et al., incorporated herein by reference, describes exemplary chelates. Examples include nitrilotriacetic acid, 2,2'-bis(salicylideneamino)-6,6'-demethyldiphenyl, and 1,8-bis(a-pyridyl)-3,6-dithiaoctane, or the like.

"Signaling entity" means an entity that is capable of indicating its existence in a

25   particular sample or at a particular location. Signaling entities of the invention can be those that are identifiable by the unaided human eye, those that may be invisible in isolation but may be detectable by the unaided human eye if in sufficient quantity (e.g., colloid particles), entities that absorb or emit electromagnetic radiation at a level or within a wavelength range such that they can be readily detected visibly (unaided or

30   with a microscope including an electron microscope or the like), or spectroscopically, electroactive entities that can be detected electronically or electrochemically, such as redox-active molecules exhibiting a characteristic oxidation/reduction pattern upon

exposure to appropriate activation energy ("electronic signaling entities"), or the like. Examples include optically active entities such as dyes, pigments, transition metal complexes, redox-active metal complexes, fluorescent or phosphorescent moieties (including, by definition, fluorescent or phosphorescent proteins such as green

5  fluorescent protein (GFP), phosphorescent moieties), up-regulating phosphors, chemiluminescent entities, electrochemiluminescent entities, or enzyme-linked signaling moieties including horse radish peroxidase and alkaline phosphatase. "Precursors of signaling entities" are entities that by themselves may not have signaling capability but, upon chemical, electrochemical, electrical, magnetic, or physical

10  interaction with another species, become signaling entities. An example includes a chromophore having the ability to emit radiation within a particular, detectable wavelength only upon chemical interaction with another molecule. Precursors of signaling entities are distinguishable from, but are included within the definition of, "signaling entities" as used herein. Another example of a signaling entity is a particle

15  that is made up of material that possesses an inherent signaling capability, including those materials whose signaling capabilities requires excitation with external energy sources.  A preferred electroactive molecule as a signaling entity of the invention is a metallocene. Metallocenes that can operate as electroactive signaling elements in accordance with the invention are known.  One example of a particularly preferred

20  electroactive molecule is one containing a ferrocene or a ferrocene derivative group or derivative, such as ferrocenyl thiol ($C_{35}H_{24}FeS$); however, other organic complexes of transitions metals are also contemplated as signaling elements.

As used herein, "fastened to or adapted to be fastened", in the context of a species relative to another species or to a surface of an article, means that the species is

25  chemically or biochemically linked via covalent attachment, attachment via specific biological binding (e.g., biotin/streptavidin), coordinative bonding such as chelate/metal binding, or the like. For example, "fastened" in this context includes multiple chemical linkages, multiple chemical/biological linkages, etc., including, but not limited to, a binding species such as a peptide synthesized on a polystyrene bead, a

30  binding species specifically biologically coupled to an antibody which is bound to a protein such as protein A, which is covalently attached to a bead, a binding species that forms a part (via genetic engineering) of a molecule such as GST or Phage, which in

turn is specifically biologically bound to a binding partner covalently fastened to a surface (e.g., glutathione in the case of GST), etc. As another example, a moiety covalently linked to a thiol is adapted to be fastened to a gold surface since thiols bind gold covalently. Similarly, a species carrying a metal binding tag is adapted to be fastened to a surface that carries a molecule covalently attached to the surface (such as thiol/gold binding) which molecule also presents a chelate coordinating a metal. A species also is adapted to be fastened to a surface if a surface carries a particular nucleotide sequence, and the species includes a complementary nucleotide sequence.

"Covalently fastened" means fastened via nothing other than one or more covalent bonds. E.g. a species that is covalently coupled, via EDC/NHS chemistry, to a carboxylate-presenting alkyl thiol which is in turn fastened to a gold surface, is covalently fastened to that surface.

"Specifically fastened" or "adapted to be specifically fastened" means a species is chemically or biochemically linked to another specimen or to a surface as described above with respect to the definition of "fastened to or adapted to be fastened", but excluding all non-specific binding.

"Non-specific binding", as used herein, is given its ordinary meaning in the field of biochemistry.

"Colloids", as used herein, means nanoparticles, i.e. very small, self-suspendable or fluid-suspendable particles including those made of material that is, e.g., inorganic or organic, polymeric, ceramic, semiconductor, metallic (e.g. gold), non-metallic, crystalline, amorphous, or a combination. Typically, colloid particles used in accordance with the invention are of less than 250 nm cross section in any dimension, more typically less than 100 nm cross section in any dimension, and in most cases are of about 2-30 nm cross section. One class of colloids suitable for use in the invention is 10-30 nm in cross section, and another about 2-10 nm in cross section. As used herein this term includes the definition commonly used in the field of biochemistry.

A "moiety that can coordinate a metal", a used herein, means any molecule that can occupy at least two coordination sites on a metal atom, such as a metal binding tag or a chelate.

As used herein, a component that is "immobilized relative to" another component either is fastened to the other component or is indirectly fastened to the

other component, e.g., by being fastened to a third component to which the other component also is fastened, or otherwise is translationally associated with the other component. For example, a signaling entity is immobilized relative to a binding species if the signaling entity is fastened to the binding species, is fastened to a colloid particle to which the binding species is fastened, is fastened to a dendrimer or polymer to which the binding species is fastened, etc. A colloid particle is immobilized relative to another colloid particle if a species fastened to the surface of the first colloid particle attaches to an entity, and a species on the surface of the second colloid particle attaches to the same entity, where the entity can be a single entity, a complex entity of multiple species, a cell, another particle, etc. In all embodiments of the invention, where a species is described as immobilized relative to another entity (another species, a surface, etc.), it to be understood that the species can be fastened to the entity in some embodiments, where those of ordinary skill in the art would understand that it is possible for the species to be fastened to the entity.

"Diverse biological species" means different animals, such as mouse and hamster, mouse and goat, etc.

The term "sample" refers to any cell, tissue, or fluid from a biological source (a "biological sample", or any other medium, biological or non-biological, that can advantageously be evaluated in accordance with the invention including, but not limited to, a biological sample drawn from a human patient, a sample drawn from an animal, a sample drawn from food designed for human consumption, a sample including food designed for animal consumption such as livestock feed, milk, an organ donation sample, a sample of blood destined for a blood supply, a sample from a water supply, or the like. One example of a sample is a sample drawn from a human or animal to whom a candidate drug has been given to determine the efficacy of the drug.

A "sample suspected of containing" a particular component means a sample with respect to which the content of the component is unknown. For example, a fluid sample from a human suspected of having a disease, such as a neurodegenerative disease or a non-neurodegenerative disease, but not known to have the disease, defines a sample suspected of containing neurodegenerative disease aggregate-forming species. "Sample" in this context includes naturally-occurring samples, such as physiological samples from humans or other animals, samples from food, livestock feed, etc., as well

as "structurally predetermined samples", which are defined herein to mean samples, the chemical or biological sequence or structure of which is a predetermined structure used in an assay designed to test whether the structure is associated with a particular process such as a neurodegenerative disease. For example, a "structurally predetermined sample" includes a peptide sequence, random peptide sequence in a phage display library, and the like. Typical samples taken from humans or other animals include cells, blood, urine, ocular fluid, saliva, cerebro-spinal fluid, fluid or other samples from tonsils, lymph nodes, needle biopsies, etc.

As used herein, "metal binding tag/metal/chelate linkage" defines a linkage between first and second species in which a first species is immobilized relative to a metal binding tag and a second species is immobilized relative to a chelate, where the chelate coordinates a metal to which the metal binding tag is also coordinated. U.S. Patent No. 5,620,850 of Bamdad, et al., incorporated herein by reference, describes exemplary linkages.

The term "biological binding" refers to the interaction between a corresponding pair of molecules that exhibit mutual affinity or binding capacity, typically specific or non-specific binding or interaction, including biochemical, physiological, and/or pharmaceutical interactions. Biological binding defines a type of interaction that occurs between pairs of molecules including proteins, nucleic acids, glycoproteins, carbohydrates, hormones and the like. Specific examples include antibody/antigen, antibody/hapten, enzyme/substrate, enzyme/inhibitor, enzyme/cofactor, binding protein/substrate, carrier protein/substrate, lectin/carbohydrate, receptor/hormone, receptor/effector, complementary strands of nucleic acid, protein/nucleic acid repressor/inducer, ligand/cell surface receptor, virus/ligand, etc.

The term "binding partner" refers to a molecule that can undergo binding with a particular molecule. Members of pairs of molecules that can undergo biological binding, as exemplified above, are examples. For example, Protein A is a binding partner of the biological molecule IgG, and vice versa.

The term "determining" refers to quantitative or qualitative analysis of a species via, for example, spectroscopy, ellipsometry, piezoelectric measurement, immunoassay, electrochemical measurement, and the like. "Determining" also means detecting or quantifying interaction between species, e.g. detection of binding between two species.

The term "self-assembled monolayer" (SAM) refers to a relatively ordered assembly of molecules spontaneously chemisorbed on a surface, in which the molecules are oriented approximately parallel to each other and roughly perpendicular to the surface. Each of the molecules includes a functional group that adheres to the surface, and a portion that interacts with neighboring molecules in the monolayer to form the relatively ordered array. A wide variety of SAMs can be used in accordance with the invention, on a wide variety of surfaces, to present desired species such as binding partners, signaling entities, and the like at a surface of an article such as an electrode, colloid particle, or the like. Those of ordinary skill in the art can select from among a wide variety of surfaces, functional groups, spacer moieties, etc. for forming SAMs. An exemplary description can be found in U.S. Patent No. 5,620,850. See also Laibinis, P. E.; Hickman, J.; Wrighton, M. S.; Whitesides, G. M. Science 245, 845 (1989), Bain, C.; Evall, J.; Whitesides, G. M. J. Am. Chem. Soc. 111, 7155-7164 (1989), Bain, C.; Whitesides, G. M. J. Am. Chem. Soc. 111, 7164-7175 (1989), each of which is incorporated herein by reference. The formation of SAMs on fluid-suspendable particles such as colloid particles is described in U.S. Patent Application Serial No. 09/602,778, filed June 23, 2000, entitled "Interaction of Colloid Immobilized Species with on Non-Colloidal Structures", by Bamdad, et al., incorporated herein by reference. Certain embodiments of the invention make use of self-assembled monolayers (SAMs) on surfaces, such as surfaces of colloid particles, and articles such as colloid particles having surfaces coated with SAMs. In one set of preferred embodiments, SAMs formed completely of synthetic molecules completely cover a surface or a region of a surface, e.g. completely cover the surface of a colloid particle. "Synthetic molecule", in this context, means a molecule that is not naturally occurring, rather, one synthesized under the direction of human or human-created or human-directed control. "Completely cover" in this context, means that there is no portion of the surface or region that directly contacts a protein, antibody, or other species that prevents complete, direct coverage with the SAM. I.e. in preferred embodiments the surface or region includes, across its entirety, a SAM consisting completely of non-naturally-occurring molecules (i.e. synthetic molecules). The SAM can be made up completely of SAM-forming species that form close-packed SAMs at surfaces, or these species in combination with molecular wires or other species able to promote electronic

- 16 -

communication through the SAM (including defect-promoting species able to participate in a SAM), or other species able to participate in a SAM, and any combination of these. Preferably, all of the species that participate in the SAM include a functionality that binds, optionally covalently, to the surface, such as a thiol which

5      will bind to a gold surface covalently. A self-assembled monolayer on a surface, in accordance with the invention, can be comprised of a mixture of species (e.g. thiol species when gold is the surface) that can present (expose) essentially any chemical or biological functionality. For example, they can include tri-ethylene glycol-terminated species (e.g. tri-ethylene glycol-terminated thiols) to resist non-specific adsorption, and

10     other species (e.g. thiols) terminating in a binding partner of an affinity tag, e.g. terminating in a chelate that can coordinate a metal such as nitrilotriacetic acid which, when in complex with nickel atoms, captures a metal binding tagged-species such as a histidine-tagged binding species. The present invention provides a method for rigorously controlling the concentration of essentially any chemical or biological

15     species presented on a colloid surface or any other surface. Without this rigorous control over peptide density on each colloid particle, co-immobilized peptides would readily aggregate with each other to form micro-hydrophobic-domains that would catalyze colloid-colloid aggregation in the absence of aggregate-forming species present in a sample. This is an advantage of the present invention, over existing colloid

20     agglutination assays. In many embodiments of the invention the self-assembled monolayer is formed on gold colloid particles. Self-assembled monolayers can be made to be electrically conductive. As a working example, fig. 16 shows ACV demonstration of enhanced electronic communication across a self-assembled monolayer, and redox signaling of protein immobilization to a cell surface, against a

25     control.

A "self-assembled monolayer-forming species" comprises a species that, when exposed to an appropriate surface with other, like species, e.g. provided with like species in an appropriate solution and exposed to an appropriate surface, will spontaneously form a self-assembled monolayer on the surface.

30     A species "able to integrate into a self-assembled monolayer" (which can be a self-assembled monolayer forming species) is a species having a chemical functionality favoring participation in a self-assembled monolayer comprising the species and other,
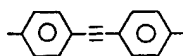
- 17 -

self-assembled monolayer-forming species with which it is not chemically
incompatible. For example, the species may include a functional group selected to
adhere to a surface on which the self-assembled monolayer is formed, and may include
a remainder portion that may be approximately linear (not highly-branched), but which
does not facilitate close packing. Molecules including a significant amount of
unsaturation, for example a series of interconnected aromatic rings, are examples. Such
a species may or may not be a self-assembled monolayer-forming species. Typically,
species that are able to integrate into a self-assembled monolayer but are not able
themselves to form a self-assembled monolayer will be able to participate in formation
of and integrate into a self-assembled monolayer when present in an amount of up to
about 50% as a percentage of overall species including chemically-compatible self-
assembled monolayer-forming species.

The term "self-assembled mixed monolayer" refers to a heterogeneous self-
assembled monolayer, that is, one made up of a relatively ordered assembly of at least
two different molecules.

"Molecular wires" as used herein, means wires that enhance the ability for a
fluid encountering a SAM-coated electrode to communicate electrically with the
electrode. This includes electrically conductive molecules or molecules that can cause
defects in the SAM allowing communication with the electrode. A non-limiting list of
additional molecular wires includes 2-mercaptopyridine, 2-mercaptobenzothiazole,
dithiothreitol, 1, 2-benzenedithiol, 1, 2-benzenedimethanethiol, benzene-ethanethiol,
and 2-mercaptoethylether. Conductivity of a monolayer can also be enhanced by the
addition of molecules that promote conductivity in the plane of the electrode.
Conducting SAMs can be composed of, but are not limited to: 1) poly (ethynylphenyl)
chains terminated with a sulfur; 2) an alkyl thiol terminated with a benzene ring; 3) an
alkyl thiol terminated with a DNA base; 4) any sulfur terminated species that packs
poorly into a monolayer; 5) all of the above plus or minus alkyl thiol spacer molecules
terminated with either ethylene glycol units or methyl groups to inhibit non specific
adsorption. A variety of molecules can be used for this purpose, including but not
limited to poly (ethynylphenyl thiol) (i.e. $C_{16}H_{10}S$), referred to herein as MF1.:



SH

- 18 -

Thiols are described because of their affinity for gold in ready formation of a SAM. Other molecules can be substituted for thiols as known in the art from U.S. Patent No. 5,620,820, and other references. Molecular wires typically conduct electronically or, because of their bulk or other conformation, creates defects in an otherwise relatively

5      tightly-packed SAM to prevent the SAM from tightly sealing the surface against fluids to which it is exposed. The molecular wire causes disruption of the tightly-packed self-assembled structure, thereby defining defects that allow fluid to which the surface is exposed to communicate electrically with the surface. In this context, the fluid communicates electrically with the surface by contacting the surface or coming in close

10     enough proximity to the surface that electronic communication via tunneling or the like, can occur.

A "chimeric oligo solution" is an oligonucleotide sequence, such as DNA, that is simultaneously complimentary to two oligonucleotide identifiers each of which corresponds to two different binding partners. A complete set of chimeric oligo

15     solutions represents a set of all possible combinations of interacting binding partners in any given procedure.

The present invention provides methods by which a large number of proteins, such as those encoded by entire genomes of a species, can be simultaneously tested for interaction with any known or unknown component of a genome or with any chemical

20     species. The invention also provides methods for detecting interactions involving proteins in which the nucleic acids, which encode them, are immediately available once the protein has been selected as an interacting species.

The invention also provides methods for detecting interactions between genetically encoded species and chemical species in order to identify new affinity

25     reagents for biological and biochemical studies. Much of the following description involves a variety of methods, compositions and species, and articles for monitoring (detecting) interactions between chemical or biological species including techniques useful for drug screening. Major features of the following aspects of the invention include the following. Tools for proteomic studies including protein chips and particles

30     for signaling interactions, and multi-particle systems such as two-particle systems. In multi-particle systems one particle can be a recruitable particle and the other particle can carry a binding partner of an agent presented by the recruitable particle and can

- 19 -

also be a signaling entity or carry an auxiliary signaling entity. Another major area involves cell studies, especially techniques involving interactions between ligands and cell surface proteins and receptors. Discovery and therapeutics involving drugs that can effect these interactions also is described, with an emphasis on drug therapy

5      involving angiogenesis. Specifically, cell receptor/ligand interactions that can inhibit or promote angiogenesis are described. Another area involves detecting proteins, either in solution or on the surfaces of intact cells, for diagnostic purposes.

A major disadvantage of existing *in vitro* binding assays is that they are not compatible with high throughput. Proteins must be sequentially tested in pair-wise

10     binding assays because: 1) proteins cannot be amplified as nucleic acids can, making identification by sequencing after a binding assay difficult or impossible; therefore, the identity of each putative binding partner must be carefully tracked; 2) typical binding assays produce a single type of signal so that each pair to be tested must be kept in isolation so that a positive signal can be assigned to the appropriate binding partners.

15     Among other aspects, the present invention solves these problems by providing a convenient method of tracking proteins that have been pooled, through the use of coding tags (identifiers) that are linked to putative binding species, which include proteins as well as chemical species. Methods of the invention detail techniques for "connecting" a coding identifier to a biological or chemical species via co-

20     immobilization on a common surface, which in a preferred embodiment is the surface of a particle. In another preferred embodiment, an oligo is used to identify the biological or chemical species, wherein the 4-bit code of a DNA uniquely identifies the co-immobilized species. In an especially preferred embodiment, the biological species under study is a protein that is expressed off of the encoding DNA, which uniquely

25     identifies it. The invention also describes methods that facilitate the attachment of the identifier to a surface to which the biological or chemical species is also attached. In a preferred embodiment, an expressed protein and its encoding plasmid DNA are attached to a common particle via an affinity tag on the protein binding to a metal chelate on the particle and a DNA recognition motif contained within the plasmid

30     binding to DNA-binding proteins on the particle surface. The invention further describes high throughput methods for detecting and selecting interacting partners, then rapidly identifying the interacting partners.

- 20 -

A variety of techniques and components associated with various assays, kits, detection methods, etc. are described below. It is to be understood that the techniques of the invention involving oligonucleotide identifiers and proteins can be used in conjunction with any of the specific assays described herein, and these assays are

5      provided by way of example only, as oligonucleotide identifier techniques of the invention can be used in conjunction with essentially any biological or chemical binding assay. Oligonucleotide identifier techniques of the invention are particularly well-suited to assays involving particles, beads, chips and colloids, in order to rapidly identify interacting protein partners from a pool of putative interactors, which are

10     described below.

Those of ordinary skill in the art will clearly understand where, in the following description, oligonucleotide identifier techniques of the invention can be used, and where oligonucleotide identifier techniques of the invention can be used in essentially any assay technique.

15     In one aspect, the present invention contemplates interaction between chemical or biological agents for analysis, drug screening, or the like. The invention includes but is not limited to analyzing and/or inhibiting protein-protein interactions, protein-chemical species interactions, ligand-nucleic acid interactions, ligand-receptor interactions, including but not limited to ligands on intact cells (growing on an

20     electrode, or in solution or in suspension). The present invention contemplates a variety of embodiments, including the use of drug candidates, known or putative ligands, and small molecule drug libraries.

One aspect of the invention involves oligonucleotide identifiers, which by definition include any number of bases (nucleotides), in which the 4-bit nucleic acid

25     code is used to form a sequence that uniquely identifies some natural or synthetic material. This includes natural or synthetic nucleotide sequences, or derivatives of nucleic acids (including DNA and thiol-modified DNA, nucleotides fastened to polymer backbones, etc) that are adapted to be fastened to surfaces, that also can carry potential chemical or biological binding partners. The oligonucleotide identifiers can

30     be short DNA sequences, for example from about 2 to about 20 bases in length, preferably from about 6 to about 12 bases in length. Longer oligonucleotide identifiers

can be used as well, for example those of up to 50, or 100, or several hundred bases in length.

According to the invention, an oligonucleotide identifier is attached to a surface to which a corresponding chemical or biological agent also is attached and which it will

5    uniquely identify. The surface can be essentially any surface useful in chemical or biological analysis, including all surfaces described above, such as surfaces of particles such as fluid-suspendable particles or non-suspendable particles, larger surfaces such as those of chips, microarray chips, surfaces involved in electronic detection assays, cell surfaces etc. For example, the surface may be the surface of a colloid, where each

10   colloid or set of colloids displays a single binding species and a single oligonucleotide identifier. Alternatively, the surface may be the surface of a spatially addressable array chip where multiple pairs of agents and identifiers are fastened in relatively close proximity to each other, but separated to the extent that they can be individually formed and analyzed. Each distinct spatial address of the chip displays a potential binding

15   partner and an identifying oligonucleotide identifier nearby.
Species of interest, which are surface-immobilized relative to their unique encoding identifiers, are allowed to interact with other species in solution or attached to other surfaces or particles. Interacting species are isolated using any of a variety of techniques. The identities of the interacting species are then rapidly determined by

20   sequencing, hybridizing, or otherwise determining the sequence of the attached identifier.

Surfaces to which oligonucleotide identifiers of the invention are fastened or adapted to be fastened also can carry immobilized signaling entities such as those described herein and in International Pat. Apl Ser. No: PCT/US01/40801, filed

25   05/25/01, entitled, "Electroactive Surface-Confinable Molecules", by Bamdad, et al., incorporated herein by reference, as well as other documents incorporated herein by reference. As will be apparent from the description below, particles such as colloid particles can participate in oligonucleotide-identified interactions and may or may not carry auxiliary signaling entities. For example, in one set of embodiments colloid

30   particles are used that are free of auxiliary signaling entities where the colloid particles themselves serve as signaling entities via color change upon agglomeration. In another set of embodiments colloid particles or other surfaces carry auxiliary signaling entities

- 22 -

such as fluorescent markers (optionally different fluorescent markers at different wavelengths for different particles), electroactive species such as ferrocenes (optionally different ferrocenes with different oxidation/reduction potentials on different articles), etc.

5          Fastening of oligonucleotide identifiers and chemical or biological species that may be binding partners to surfaces can be carried out according to any technique known in the art. Preferred techniques involve the use of self-assembled monolayers on surfaces. Self-assembled monolayer-forming species, or species able to integrate into a self-assembled monolayers can include chemical or biological species to be

10         studied or oligonucleotide identifiers, and can thereby be incorporated into SAMs. Alternatively, species able to form or integrate into SAMs can include linkers for attachment to chemical or biological species or oligonucleotide identifiers, and the oligonucleotide can be attached to the surface after SAM formation via the linkers. Such linkers can include affinity tags or species that bind to affinity tags, species

15         suitable for EDC/NHS coupling, oligonucleotide linkers, biotin-streptavidin interaction, species that can participate in DNA ligation techniques etc. Examples of DNA ligation techniques include incorporation of a specific oligonucleotide sequence into a SAM, that encodes for a restriction site, and ligating a second oligonucleotide sequence, terminated in the same restriction site, to the original oligonucleotide sequence. An

20         example of a commonly used ligase is T4 Ligase. Alternatively, a blunt-end oligonucleotide could be added using blunt-end ligation techniques. Preferred species, for use with affinity tags include metals coordinated by chelates, for use with polyamino acid tags. For example, surfaces can be coated with SAMs exposing metals coordinated by chelates, and chemical or biological species, or oligonucleotide

25         identifiers, can carry polyamino acid tags for coordination to the metal thereby linking the chemical or biological species, or oligonucleotide identifier, to the surface. Different chemistries, or the same chemistry can be used for linkage of any species or identifiers involved in the invention to a single surface.

            Referring now to Fig. 1, a surface of a colloid particle 140 adapted to bind

30         essentially any chemical or biological species and also to bind an oligonucleotide identifier is illustrated schematically. Colloid 140 includes a surface 142 upon which is a SAM 144. SAM 144 is only partially illustrated - the SAM preferably will

completely coat the surface 142. SAM 144 includes one species 146 that exposes, away from surface 142, terminating in a chelate able to coordinate a metal, or a chelate coordinating a metal, 148 (represented in the figure as NTA, nitrilotriacetic acid). A chemical or biological species 154, comprising a polyamino acid tag 156, when exposed to colloid 140, will fasten to chelate/metal 148.

SAM 144 also includes a species 150 comprising an oligonucleotide linker 152. An oligonucleotide species 158, including a linker portion 160 that is the complement of linker 152, and a section 162 defining an oligonucleotide identifier, when exposed to colloid particle 140 will fasten thereto. Thus, a plurality of colloid particles 140, each including an identically-derivatized surface can be provided as a kit. In an assay, different chemical or biological species 154 can be attached to one set of colloid particles to which is attached a unique oligonucleotide identifier 162. Different batches of colloid particles can carry different chemical or biological species and corresponding oligonucleotide identifiers. Of course, colloid particles can include solely species 146 where chemical or biological species and oligonucleotide identifiers each include polyamino acid tags 156, or colloid particles can include solely species 150 where both the chemical or biological species, and the oligonucleotide identifier, each includes an oligonucleotide linker 160. That is, oligonucleotide identifiers that uniquely identify a species attached to a surface need not be directly attached to the common surface. For example, for convenience, surfaces may be derivatized with a universal oligonucleotide. A second oligonucleotide comprised of a portion that is complementary to the universal DNA sequence, which is directly attached to the surface, and a second portion which uniquely identifies a chemical or biological species that is also attached to the surface.

The arrangement of Fig. 1 can be used to provide a set of colloid particles is provided, each carrying an immobilized chemical or biological species such as a biological binding partner (e.g., a protein or small molecule), and each an immobilized oligonucleotide identifier. A record is made of the sequence of the identifier, and the identity of the chemical or biological species (potential binding partner) immobilized to the same colloid particles as the identifier. As an example of procedure, sets of colloids each bearing a distinct species are pooled together in solution and allowed to interact. This next step is an alternative to separating out interacting particle-immobilized

- 24 -

species, then sequentially releasing and sequencing the attached oligo identifiers in order to identify interacting species. A set of oligos is added that we call "chimeric oligo solutions". Each DNA strand in this set is comprised of a chimeric sequence that is complementary to two oligonucleotide identifiers. The entire set of "chimeric oligo

5       solutions" would contain chimeric sequences that represent all possible solutions to the problem of which species interact with each other. These "chimeric oligo solutions" are then incubated with the colloids that present putative binding partners and attached oligo identifiers. The "chimeric oligo solutions" that have simultaneously hybridized to oligonucleotide identifiers on two sets of colloids that bear interacting partners, identify

10      which putative binding species interact with each other. The hybridized "chimeric oligo solutions" need to be separated from oligos that remain free in solution as well as from oligos that have hybridized to only one sequence identifier. Free oligos are easily removed by pelleting the colloids and discarding the supernatant. Chimeric oligo solutions that are hybridized to only one sequence identifier can be distinguished from

15      those hybridized to two enzymatically digesting the free end with enzymes that degrade single stranded nucleic acid strands. Chimeric oligosolutions are then released from the particles by any one of a number of methods including dissociation by heated water, chemical release, etc. The chimeric oligosolutions are then sequenced to reveal the identity of the interacting partners. The chimeric oligosolutions can also be

20      enzymatically amplified (such as by PCR) prior to the sequencing step.

In one embodiment, the invention involves an oligonucleotide identifier that uniquely identifies a protein wherein the oligonucleotide identifier is the very sequence that encodes the protein. Both the expressed protein and the oligonucleotide identifier (e.g. DNA) that encoded it are immobilized relative to each other, e.g. each is

25      immobilized relative to a linker species, so that the protein is presented for binding studies and its oligonucleotide identifier is retrievable after selection of interacting pairs of proteins.

The oligonucleotide identifier can be in essentially any form, e.g. plasmid form, as in a protein expression vector, or in a linear form, such as nucleic acids strands that

30      are generated by a polymerase chain reaction (PCR). PCR generated fragments can be engineered to include sequences, such as a start site of transcription, to promote protein expression then translation. These nucleic acid templates are especially well suited for

- 25 -

*in vitro*, cell-free protein expression systems such as the Rapid Translation System (RTS) sold by Roche Diagnostics.

This embodiment of the invention is facilitated as follows. A protein is expressed from an expression vector or template. The solution in which the protein is expressed will then contain both the expressed protein of interest (which can be a putative binding partner) and the DNA that encoded the protein (which will serve as the oligonucleotide identifier). Both the protein of interest and the oligonucleotide identifier that encodes the protein can be immobilized to a common linker species by a variety of methods. Where the linker species is an article, the oligonucleotide identifier can be immobilized relative to the surface of the article by a variety of methods as described herein. E.g., the surface can be derivatized to present both a moiety to facilitate the attachment of the protein of interest as well as a moiety to facilitate the attachment of the oligonucleotide identifier, as described herein. The protein can be expressed with an affinity tag from a DNA template that contains a convenient functionality that facilitates the DNA's attachment to a surface that also presents binding partners of the protein's affinity tag. For example, an affinity tagged protein can be expressed off a DNA template that bears biotin. When exposed to a surface bearing both a metal binding tag/metal/chelate linkage and streptavidin, both gene (a section of the oligonucleotide identifier) and gene product (protein) will be captured and presented on the common surface. A convenient approach is to express a histidine-tagged protein from a biotinylated template. Surfaces, including particles, such as colloids, are coated with heterologous SAMs bearing both NTA-nickel (the binding partner of the histidine tag) and biotin. The surfaces are first exposed to streptavidin, which has four binding sites for biotin. This surface will then capture and present the histidine tagged gene product as well as its encoding DNA sequence. As another example, proteins of interest fused to glutathione-S-transferase (GST) can be attached to entities or articles presenting glutathione. Alternatively, the protein of interest can be expressed as a fusion protein with thioredoxin, then attached to a surface that presents a binding partner of thioredoxin. GST and thioredoxin serve a dual purpose. First, these proteins provide a convenient affinity tag for the protein of interest. Secondly, proteins commonly used as fusion partners increase the solubility of the

expressed protein of interest and thus increase the effective concentration of the protein.

In another technique for forming an oligonucleotide identifier immobilized relative to a protein for which is encodes, via a linker species, the oligonucleotide
5    identifier may also contain a binding site for a DNA/RNA-binding protein. This sequence is preferably inserted downstream of the gene to be expressed. The protein that binds to the binding site on the oligonucleotide identifier can then be immobilized at a surface (e.g. on a particle) which also presents a moiety for the attachment of the expressed protein (e.g. NTA/nickel). In one embodiment, a nucleic acid binding
10   protein is immobilized on the surface and used to capture the oligonucleotide identifier. Nucleic acid binding proteins bind to DNA or RNA by recognizing either a specific nucleic acid sequence motif or by recognizing a tertiary structure. For example, a common DNA-binding yeast protein is Gal4. It binds as a dimer to a specific sequence of double stranded DNA, while single stranded binding protein binds to single stranded
15   DNA. Other proteins bind to structural elements of DNA or RNA, such as binding to cruciform DNA, hairpins or to specific RNA loops.

Ideally, the protein binding site sequences inserted into the template that encodes the protein of interest are recognition motifs for a protein from a species distinct from that of the expression system. For example, if the protein is expressed
20   using an E. *coli* based system, then the oligonucleotide identifier is attached to the common surface by binding to a yeast protein, such as Gal4. In this way, the protein expression system does not include or produce extraneous proteins that would compete for the cognate binding site on the oligonucleotide identifier.

In an alternative approach, the protein of interest is expressed (by an
25   oligonucleotide identifier) as a fusion protein with a protein fragment that has a binding partner that can be attached to the oligonucleotide identifier. This provides a method for attaching the oligonucleotide identifier to the protein of interest. There are a variety of proteins or fragments of proteins that can be genetically fused to the protein of interest, which also bind to small molecules which can be used to modify DNA. Short
30   strands of DNA that are so modified can be used as primers in PCR reactions to produce expression templates from which fusion proteins that capture their own oligonucleotide identifier can be produced.

- 27 -

For example, a PCR product that contains sequences that encode the protein of interest as well as sequences that encode streptavidin can be generated using biotinylated primers. The resultant PCR product, which will also include elements necessary for transcription/translation, serves as a template for protein expression and also yields an oligonucleotide identifier that has been adapted to be fastened to the protein expressed via the biotin-streptavidin linkage. If desired, the streptavidin may be expressed as a fragment that only binds one biotin. Additionally, the fusion protein which contains streptavidin and the protein of interest may also contain an affinity tag to facilitate attachment of the protein/oligonucleotide identifier to a common surface or particle.

In another example, the protein of interest is expressed as a fusion protein with a DNA-binding protein, such as LexA, from an expression vector or template that contains DNA-binding sites for LexA. In this way, the gene product and gene are linked to each other.

To facilitate the detection of binding events between proteins of interest, one of the proteins can be modified with a signaling entity. One way to accomplish this is by direct attachment of the signaling entity to one of the proteins of interest. Another way is to genetically fuse the protein of interest to a protein that has a signaling capability. For example, a protein of interest can be expressed as a green fluorescent protein (GFP) fusion protein.

Alternatively, the proteins of interest have signaling capability when they are co-immobilized with a signaling entity on a common surface. For example, a protein of interest can be attached to a polymer to which a signaling entity such as Ru complex has also been attached. Binding partners are attached to recruitable particles such as magnetic beads and then detected by ECL methods. Similarly, proteins of interest can be co-immobilized on particles that also present signaling entities. For example a protein of interest can be attached to a particle or colloid that also presents signaling entities such as fluorescent or phosphorescent or redox-active moieties to facilitate optical or electrochemical detection, respectively.

In yet another embodiment, the oligonucleotide identifier that encodes the protein of interest is generated such that it also possesses signaling capabilities. For example, the DNA expression template can be generated by PCR using primers in

which at least some of the bases have been chemically modified with redox-active moieties such as ferrocene derivatives. The protein of interest and the oligonucleotide identifier (expression template) are then attached to a common surface, such as a colloid particle. The particles are mixed with putative binding partners attached to

5      magnetic beads, and magnetically drawn to a sensing electrode where the presence of the redox-active molecule, hence the interaction, is detected. Alternatively, bases modified with a signaling functionality can be used in the polymerase chain reaction to generate an expression template that can signal. Similarly, the protein of interest is expressed from a DNA template that has been modified with a fluorescent or

10     phosphorescent moiety.

In one example of synthesis and use, a nucleic acid is generated that encodes a histidine-tagged protein of interest (for attachment of the protein to a surface that displays an NTA-Ni$^{++}$ moiety) and that also contains a Gal4 binding site for attachment of the oligonucleotide identifier to the common surface. The protein is preferably

15     expressed in a cell-free system to minimize levels of background DNA and proteins. After protein expression, the reaction mixture is incubated with particles (e.g. colloids or beads) that present metal/chelate (e.g. NTA-Ni) moieties and Gal4 for fastening of the protein and oligonucleotide identifier, respectively. The particle-immobilized NTA-Ni$^{++}$ moieties capture and present the protein of interest (via the binding tag) and

20     the particle-immobilized Gal4 captures the DNA that encodes the expressed protein, because it also contained Gal4 recognition sites. In this way an identifying oligonucleotide identifier has been immobilized relative to the expressed protein.

These particles can then be used in a variety of assay formats. The following assay is compatible with detecting protein-protein interactions involving large numbers

25     of possible binding partners. A portion of the putative binding partners is attached to a set of colloids, which also bear Gal4 for the attachment of the nucleic acid, which encoded the protein. The colloids also carry signaling entities. The signaling entities can be electroactive, electrochemiluminescent, optical, etc. A second set of putative binding partners is attached to magnetic particles. Large numbers of magnetic particles

30     and colloidal particles are mixed together to allow binding to occur between proteins immobilized relative to the beads and particles. A magnetic field is then used to collect

the magnetic particles at a sensing location where the signals carried on the colloids can be detected.

As described, a gene product (protein) and the gene (oligonucleotide identifier) are connected to each other via attachment to a linker species. Protein recognition motifs (DNA or RNA sequences) are inserted into an expression (or translation) vector up- or downstream of the sequences that encode the gene of interest. These DNA recognition motifs facilitate attachment of the gene to a surface or particle presenting the cognate DNA-binding protein. The surface also presents a moiety to facilitate attachment of the gene product. For example, surfaces that present both NTA-Ni and Gal4 will capture histidine-tagged proteins and DNA that bears Gal4 recognition sequences (motifs). The nucleic acids referenced here can be any nucleic acid encoding the gene of interest and having the ability to promote transcription or translation, including: a protein expression vector (plasmid); a linear piece of double stranded DNA, such as a PCR product, including those designed for *in vitro* protein expression or translation.

The nucleic acid that encodes the gene of interest is chemically modified to facilitate its attachment to a surface or a particle. For example, a protein is expressed in a cell-free expression system off of a biotinylated PCR product. If the gene of interest has been modified with an affinity tag such as GST, both the gene product and the gene can be attached to a surface that presents both glutathione and streptavidin (or other avidin derivative).

The gene of interest is expressed as a fusion protein with a protein that can be readily attached to the encoding gene. For example, the gene of interest is expressed as a LexA fusion protein and the DNA template or plasmid, used for protein expression or translation, also contains DNA binding sites for LexA. As another example, the gene of interest is expressed as a streptavidin fusion protein and the DNA template used for protein expression or translation, is biotinylated.

The present invention allows the investigation of interactions between chemical or biological species, such as protein-protein binding interactions, on a very large scale. Large numbers of potential binding interactions can be investigated in a single solution. Once many potential binding partners are brought together, they must be separated to identify which were involved in binding interactions. The following is a description of

- 30 -

techniques for separating species that have been involved in interaction from species that have not.

As one example, magnetic *in situ* selection/dilution can be used, as described in a U.S. patent application, filed 10/03/01, entitled "Magnetic In Situ Dilution", by

5    Bamdad, incorporated herein by reference.

The magnetic *in situ* selection/dilution technique can be used in connection with any format in which a potential binding partner is immobilized relative to a particle that can be drawn to a magnet, where once the particle is drawn to the magnet, whether the species has been involved in a binding interaction can be detected. For example,

10   potential binding partners can be immobilized relative to magnetic particles, and species that might bind with these species can be labeled with signaling entities such that once a particle is drawn to a magnet the presence or absence of the signaling entity in proximity of the magnet can be detected. For example, as described in the above-noted patent application incorporated by reference, an array of electrodes can be

15   provided on a surface, each electrode individually addressable, and each electrode accompanied by a corresponding individually-addressable magnet. A plurality of magnetic beads can be provided, each carrying an immobilized potential biding partner. A plurality of species that are putative binding partners of the binding partners immobilized on the magnetic beads can be provided, each of these species immobilized

20   relative to a signaling entity. The binding partners on the beads can be the same or different, and the species that are their putative binding partners can be the same or different. Mixing the magnetic bead-immobilized binding partners with the species that are their putative binding partners may result in some magnetic beads immobilized relative to signaling entities (where binding has occurred between putative partners)

25   and other magnetic beads that are not immobilized relative to signaling entities. Following mixture, within a fluid medium, of the putative binding partners, and optionally without any wash step, the solution can be exposed to the electrode array. Beads can be magnetically drawn to electrodes and determination can be made as to which electrodes have not drawn beads immobilized to signaling entities (for example,

30   optically where the signaling entities are optical signaling entities or electrochemically where the electrodes corresponding to the magnetic beads generate a signal determinative of whether a signaling entity is or is not in proximity). Following this,

individual magnetic regions that have not attracted beads carrying immobilized
signaling entities are de-magnetized, releasing the beads, which are washed away.
Magnets at which beads carrying signaling entities have been attracted remain
magnetized during this washing and removal step. Subsequent to removal, all magnets

5      are released and the process is continued until, statistically, only one magnetic bead
exists for each magnet. At this point, after a magnetic attraction step, each magnet that
has not attracted a bead carrying an immobilized signaling entity is released, and all
remaining magnets will have attracted only magnetic beads carrying a pair of binding
partners indicated by the presence of a signaling entity at the magnet. At this point,

10     these species are released and the identity of binding partners is identified via
identification of oligonucleotide identifiers described herein. This is the technique
described generally in the above-identified U.S. patent applications.

       As a specific example, a first set of gene products (proteins) and their respective
genes (oligonucleotide identifiers) are attached to magnetic beads. A second set of

15     gene products and their respective genes are attached to particles, such as colloids, that
also bear electronic signaling moieties. The two particle populations are incubated
together in solution and binding interactions between proteins immobilized on different
particles are allowed to occur. When an interaction takes place between a protein on a
magnetic bead and a protein on a colloid bearing electroactive groups, the recruitable

20     particle (magnetic bead) becomes connected to the signaling particle. A magnetic field
draws the complex to a sensing electrode where an electronic signal is transduced. The
sensing electrode is an array of electrodes configured such that the recruitment of
magnetic beads to each electrode pad is individually controlled. This can be
accomplished by interfacing each electrode pad with an individually controllable

25     electromagnet. That is, an array of individually-addressable electromagnets is
fabricated (using techniques known to those of ordinary skill in the art). Where the
electronic signaling entity is a redox-active molecule such as ferrocene, an electrode is
associated with each magnet such that when beads are drawn to a magnet, the electrode
associated with the magnet is cycled and if a redox-active signaling entity is

30     immobilized relative to the bead (via binding partner interaction), then the redox
signature of the signaling entity will be detected.

- 32 -

Another example involves optical selection. In this technique, binding partners identified by immobilized oligonucleotide identifiers, that participate in binding interactions are optically detected by any technique described herein such as visual or automatic observation of color change upon colloid-colloid aggregation, colloid

5    "decoration" of beads upon aggregation, etc. As a specific example, interacting partners, attached to different particle types, can be optically selected by a variety of methods. A first set of gene products (proteins that are putative binding partners) and genes (oligonucleotide identifiers) can be attached to non-magnetic beads. A second set of gene products can be directly or indirectly modified with fluorescent or

10   phosphorescent moieties. For example, the second set of gene products can be attached to colloids that also bear fluorescent or phosphorescent moieties. The interaction between bead-immobilized and colloid-immobilized proteins causes the agglomeration of the fluorescent or phosphorescent colloids onto the larger bead and renders it (the bead) detectable and identifiable as a bead that presents a protein that is participating in

15   an interaction. In an alternative approach, the second set of gene products is expressed as a fusion protein with a DNA binding protein or domain and the is expressed or translated off of a fluorescent or phosphorescently labeled nucleic acid. As in the previous case, the interaction between bead-immobilized proteins and the fusion proteins causes the agglomeration of a fluorescent or phosphorescent species onto the

20   bead and renders it (the bead) detectable and identifiable as a bead presenting a protein that participates in an interaction.

Beads decorated with an optically detectable species can be manually isolated, prior to identifying the interacting species, by merely picking then analyzing the beads that fluoresce.

25   Alternatively, the isolation of interacting species can be automated. Beads bearing fluorescent or phosphorescent moieties can be selected and isolated by utilizing a FACS (fluorescence-activated cell sorting) system, which is a technique well-known to those of ordinary skill in the art.

Interacting particle-attached species are first isolated. After isolating interacting

30   species, the attached nucleic acid sequences are identified by sequencing, PCR, hybridization or a combination of these techniques. Alternatively, if the protein has been expressed off plasmid DNA, the encoding DNA can be released from the

particles, and plated onto growth media. Once new colonies are grown, these provide a renewable source of the encoding DNA for analysis by any of the methods mentioned above (sequencing, PCR, hybridization or combinations of these techniques).

An advantage of the present invention is that large numbers of proteins can be

5   simultaneously tested to identify interacting pairs without apriori knowledge of the identity of either protein. Using standard techniques, libraries of complementary DNA (cDNA) can be inserted into protein expression plasmids or linear nucleic acid templates that have been modified to facilitate direct or indirect attachment to the encoded protein as described herein. For example, total mammalian cDNA can be

10   inserted into plasmids that also contain sequences that encode a histidine tag and DNA binding sites for a yeast DNA binding protein, which would be presented on a particle that also presented NTA-Ni for the capture of the expressed protein. In this way, the expressed protein and its encoding DNA would be "attached" to each other via immobilization on a common particle. Similar genetic manipulations, when coupled

15   with methods of the invention designed to detect and select interacting partners, enable the characterization of large numbers of interacting proteins and elucidation of interaction networks of entire genomes.

Methods of the invention are also used to functionally characterize an unknown protein of interest. With the recent sequencing of the human genome, it is now

20   imperative to determine the function of newly identified genes. The number of genes in the human genome is estimated to be about 40,000. However, there are protein recognition motifs that are common to several proteins. One method of characterizing unknown genes is to determine which proteins the gene product interacts with. A less complex method is to determine which of the known protein recognition motifs the

25   uncharacterized protein interacts with. A library of nanoparticles can be generated that each display a different protein recognition module, such as a kinase domain, phosphorylase, a PDZ domain, GRB1 and 2 domains, ERK, kringle, WW domains and the like. Uncharacterized proteins and their encoding DNA are separately attached to magnetic beads, then mixed with a library of colloid particles each bearing an

30   interaction domain and its associated DNA. Interacting particles are then either optically or magnetically selected, then identified via analysis of attached nucleic acids or other encoding identifiers.

- 34 -

Fragments of DNA that encode proteins can be readily inserted into either plasmids or linear nucleic acid templates to facilitate expression of the encoded proteins. cDNA libraries that are inserted into a linear expression/translation template (for *in vitro* translation) offer certain advantages over insertion into plasmids. Hybrid

5    nucleic acid expression templates can also be produced by generating the cDNA library with primers that contain both sequences that encode an affinity tag for the expressed protein and a functionality to facilitate attachment of both the gene product and its encoding DNA to a common surface. Linear expression templates can be readily modified with chemical functionalities to facilitate attachment to surfaces or

10   functionalities to add a signaling capability. For example, protein encoding nucleic acid fragments can be genetically fused to nucleic acid strands that carry a biotin moiety. Biotinylated nucleic acids of any sequence are commercially available. Streptavidin has four binding sites for. Therefore, biotinylated nucleic acid expression templates bind to surfaces that present streptavidin.

15    Libraries of cDNAs that have been inserted into plasmids offer other advantages. Once an interaction has been detected, the interacting species and the plasmids that encode them can be isolated. The plasmids can then be introduced into a host cell in order to amplify the plasmid prior to sequencing or use in additional assays.

Another method of the invention involves fusing components of a cDNA

20   library, for example representing the entire mammalian genome, to nucleic acids that encode a DNA-binding protein, or fragment thereof, and the DNA sequences to which it binds. Fusion proteins expressed from these templates will be able to bind to their encoding nucleic acids.

The invention also provides methods for identifying affinity reagents for use in

25   protein purification, immobilization and other general biochemical techniques. Because the interaction aspect of the invention is not performed in a host cell, interacting components need not be entirely generated from genetic material. In this embodiment, a first set of chemical species is attached to recruitable beads along with a identifier that codes for the chemical identity of the immobilized chemical compound. A second set

30   of genetically encoded species (preferable poly amino acids 4-14 amino acids in length) is attached to a linker species that has a signaling capability. The chemical and biological components, one or both of which may be attached to surfaces, are mixed

together such that binding interactions can occur. Methods of the invention are used to select interacting components. The components of each interacting pair are then identified. This can be accomplished using methods of the invention such as sequencing attached nucleic acid identifiers or encoding plasmids or by using other

5     analytical tools, which are better suited for the identification of chemical compounds and peptides. These methods include but are not limited to mass spec techniques and peptide sequencing techniques. Methods of the invention can also be used when the surface is the surface of a chip. With reference to Fig. 2, in one embodiment, a set of derivatized colloids, each bearing a putative binding partner and identifiers, are

10    introduced to a spatially addressable array chip that presents putative binding species in close proximity to an identifying oligonucleotide identifier. Chip-immobilized species are allowed to bind to colloid-immobilized species, then colloids bearing non-binders are washed away. A set of oligos (identifiers), which contains sequences complementary to all possible pair-wise combinations of the sequences that identify

15    putative binding partners, is then incubated with the colloid-decorated chip. After rinsing, the oligos that have hybridized to the identifiers on the colloids and the chips represent identification of the putative binding species that interacted with each other. The "identifier" oligos are dissociated from the surface by any one of a number of methods including dissociation by heated water, chemical release, etc., and are then

20    sequenced to reveal the identity of the interacting partners. The identifier oligos can also be enzymatically amplified at their specific locations while on the chip (such as by PCR) prior to the sequencing step. As shown, referring to Fig. 2, a chip 164 having a surface 166 includes a plurality of spatially-addressable regions 168, 170, etc. Each region includes a chemical or biological species (putative binding species) 174, 176,

25    178, etc. Each region also includes an oligonucleotide identifier 180, 182, 184, etc. Identifiers 180, 182, and 184 uniquely identify chemical or biological species 174, 176, and 178, respectively.

Following the reaction described above, as illustrated, one colloid particle 186 remains immobilized at surface 166. Colloid 186 includes, fastened thereto, a chemical

30    or biological species 188 that binds to species 174, and oligonucleotide identifier 190. Binding of species 174 and 188 brings identifiers 180 and 190 into close proximity, whereby an interaction hybridization identifier 192 binds to the combination of

- 36 -

identifiers 180 and 190. Identification of identifier 192 identifies the sequences of
identifiers 180 and 190, identifying one or both of chemical or biological species 174
and 188.

Array chips that display a multitude of chemical or biological species that are
putative binding partners, by and unique identifying oligos nearby, can be generated by
a variety of techniques. One method involves forming heterologous self-assembled
monolayers (SAMs) on surfaces that incorporate an entity that facilitates the attachment
of a protein (e.g. a chelate/metal – terminated species that can participate in a SAM)
and a second entity that facilitates the attachment of a nucleic acid species (e.g. a
nucleic acid linker - terminated species that can participate in a SAM). For example, a
mixed SAM can be formed from mixed thiol species that include a thiol-modified
strand of DNA and a thiol terminated in nitrilo tri-acetic acid (NTA). NTA, when
complexed with nickel, selectively captures histidine-tagged proteins. This
heterologous SAM would then be able to capture any histidine-tagged protein and any
strand of DNA that contains a sequence complementary to that displayed on the chip.
Alternatively, a SAM exposing a single linking species (e.g. chelate/metal) can be
.formed, and used with both a binding partner that carries a polyamino acid tag and an
oligonucleotide modified with a polyamino acid tag. The SAM-forming step is usually
performed by incubated a gold-coated surface with various thiols in organic solvents.
Organic solvents spread on metal surfaces. However, once the SAM has been formed,
the steps of incubating a protein with the surface, then hybridizing a unique nucleic
acid strand to the surface via a common "tail" that hybridizes to the surface-
immobilized oligo, are both performed in aqueous buffer which beads up on SAM-
coated surfaces. In this way, an entire surface may be coated with a universal SAM,
which is then "dotted" with small volumes of protein and DNA to generate an array
chip.

Alternatively, various species of DNA identifiers that are able to bind to a
surface, for example via thiol modification, in aqueous solution, can be dotted onto an
underivatized gold-coated surface. After some incubation period, the chip is then
exposed to a solvent containing "filler" thiols. These filler species may also contain a
thiol species that facilitates the attachment of proteins to the surface, which may be
carried out in a spatially addressable way.

- 37 -

Standard gene chips can also be used with the methods described herein, if they are modified to allow the placement of proteins in close proximity to an identifier.

Referring now to Fig. 3, another embodiment of the invention is illustrated. A chip 240 includes a surface 242 to which a plurality of chemical or biological species 244 – 250, etc., are fastened. Exposure of the surface to colloid particles 252 each carrying a chemical or biological species 254 and an oligonucleotide identifier 256 results in binding between species 246 and 254, as illustrated. Subsequently, localized cleavage and identification of identifier 256, or localized PCR or hybridization can identify binding. Species 244 – 250 can be identical, with non-identical species attached to colloids, or species 244 – 250 can be different with identical species attached to colloids exposed to the surface, or both species attached to the surface and the species attached to the colloids can be varied.

Referring now to Fig. 4, one arrangement for fastening an oligonucleotide to a surface is illustrated. Fig. 4 illustrates a self-assembled monolayer-forming species incorporating an oligonucleotide identifier. Specifically, the self-assembled monolayer-forming species is a long-chain alkyl thiol including a restriction enzyme cleavage site 100, a DNA priming region 102 linked to the cleavage site, and a 5-base oligonucleotide identifier 104 linked to the priming region. It is to be understood that the priming region can be defined by any entity that can link the oligonucleotide identifier to the self-assembled monolayer-forming species and also promote sequencing as will be described below with reference to Fig. 5. It is also to be understood that the arrangement of Fig. 4 is by example only, and other techniques for linking an oligonucleotide identifier to a surface can be used.

Fig. 6 illustrates incorporation of the species of Fig. 4 into a self-assembled monolayer onto a surface, specifically a surface of a colloid, along with immobilization of a biological binding partner to the same surface. A colloid particle 106 has a gold surface to which thiols will bind and upon which thiol-contained self-assembled monolayer-forming species will form self-assembled monolayers. A self-assembled monolayer 108 is formed on a surface 110 of colloid 106, including both the species of Fig. 4 and an immobilized biological binding partner (protein) 112. As illustrated, species 112 is fastened to a self-assembled monolayer-forming species that is

- 38 -

incorporated into SAM 108. Specifically, the species is linked to the self-assembled monolayer-forming species via a metal binding tag/metal/chelate linkage 114.

Fig. 7 illustrates, schematically, colloid particle 106 carrying binding species 112 and a second colloid 118 carrying a chemical or biological species 120 which is a biological binding partner of species 112. As illustrated, species 112 and 120 are each protein, although as would be understood by those of ordinary skill in the art, other binding species can be used. For example, any of a variety of proteins, peptides, or other species can define species 112 and 120, and species 112 and 120 can be fastened to the colloid particle via other affinity tags. The oligonucleotide identifier and/or chemical or biological species also can be fastened to the surface of the colloid via any of a variety of affinity tags, or a carboxylic acid thiol via EDC/NHS coupling. Alternatively, a species of interest can be attached to a thiol for direct attachment to the colloid.

Self-assembled monolayer-forming species, as illustrated, can include long carbon chains such as 11 carbons or greater.

Once species 112 and 120 are allowed to participate in a binding assay (including any assay described in the documents incorporated herein by reference such as bead coloration assays, colloid-colloid assays, etc.), if one of species 112 or 120 is identified as a binding partner in the assays, the identity of which would be desirably known, then the identity of the species can be uncovered as follows. For purposes of this discussion it is assumed that colloid 106 carries a known binding partner 112 and colloid 118 carries a species 120 that was not known to be a biological binding partner of species 112. Thus, once identification of the binding between species 112 and 120 is known (for example via aggregation of colloid particles), where a variety of other colloid particles carrying species other than 120 were involved in the assays and may have bound to species 112, then the identity of species 120 is desirably determined. With reference to Fig. 8, a restriction enzyme is added to cleave oligonucleotide identifier 104 from colloid particle 118 (along with the priming region 102). Subsequently, with reference to Fig. 5 an oligonucleotide primer 122 complementary to the priming region 102 is added, and normal PCR-based sequencing, or other standard sequencing, is performed on the cleaved oligonucleotide to decipher the sequence of the oligonucleotide identifier. Standard fluorescent sequencing can be carried out, for

- 39 -

example. Once the sequence of the oligonucleotide identifier 104 is identified, it in turn identifies species 120 as that species that had been bound to colloid particle 118 to which identifier 104 also had been bound; identifier 104 had been correlated to species 120 prior to running of the assay.

5          Another technique for use of oligonucleotide identifiers is described now with reference to Figs. 9-15. Fig. 9 illustrates an oligonucleotide identifier 124 which is part of a self-assembled monolayer-forming species 125. Identifier 124 can be fastened to and form a part of SAM-forming species 125 via any technique, and need not include a restriction enzyme cleavage site or priming region 100 and 102, respectively, as

10        illustrated in the figures above. Fig. 9 also illustrates identifier 124 and a chemical or biological species 126, each fastened to the surface of a colloid particle 128 via self-assembled monolayer-forming species. That is, a SAM on the surface of colloid 128 includes both identifier 124 and species 126. Identifier 124 and species 126 can be fastened to colloid 128 in any manner known to those of ordinary skill in the art,

15        including any manner described herein. Referring now to Fig. 10, colloid 128 is brought into proximity with a second colloid particle 130, which carries its own oligonucleotide identifier 132 and its own chemical or biological species 134. If chemical or biological species 126 and 134 bind, for example if they are biological binding partners, then oligonucleotide identifiers 124 and 132 will be brought into close

20        proximity with each other. As an example of an assay shown in Fig. 10 a plurality of colloids 128 each carrying a plurality of binding partners 126 is provided and mixed with a plurality of colloids 130, each carrying a different species that may or may not be a binding partner of species 126. Colloid 130 carries species 134 which is a binding partner of species 126, as determined by the aggregation of colloid particles 128 with

25        colloid particles 130, for example. The binding of species 126 and 134 having been determined, it is desirable (in this example) to determine the identity of species 134.

          Alternatively, a large number of colloid particles carrying a wide variety of potential binding partners can be admixed, where a variety of different binding interactions may occur.

30        In any event, once binding between species 126 and 134 has occurred (e.g. via detection of a color change from pink to blue as aggregation of colloid particles occurs), then the identity of species 126, or 134, or both is determined. Referring now

to Fig. 11, identifiers 124 and 132 together define a sequence that may be
complementary to an interaction hybridization identifier 136. Where this is the case,
identifier 136 will bind to the combination of identifiers 124 and 132. This will occur
when a variety of interaction hybridization identifiers are added to the assay, each

5    corresponding to a different potential combination of oligonucleotide identifiers
fastened to colloid particles. If only two species such as proteins or small molecules
are being assayed for binding, then only one complementary sequence would need to be
added.

Subsequently, all non-bound oligonucleotide is de-activated, e.g. by adding a

10   DNAase that degrades single-stranded DNA (Fig. 12). This eliminates any
oligonucleotide identifiers present on non-interacting colloids that did not participate in
a binding event, and any auxiliary, non-bound interaction hybridization identifiers (138
and 140). Fig. 13 illustrates the result of this step.

With reference to Fig. 14, interaction hybridization identifier 136 is removed

15   from and isolated relative to identifiers 124 and 132 via, for example, denaturization
(by boiling or addition of salt or Triton solution, etc.). Then, identifier 136 is
sequenced (Fig. 15). With the identification of species 136, the identity of identifiers
124 and 132 is determined, and thereby the identity of one or both of species 126 and
134 can be determined (identifier 124 had been correlated to species 126 and identifier

20   132 had been correlated to species 134 prior to running of the assay).

Example: Demonstration of Control of SAM permeability to electrons

This example demonstrates the ability to form a SAM including enhanced
electronic communication. The SAM is formed on a surface that includes a mixture of
a first, tight-packing species and a second species of different molecular structure that

25   enhances the permeability of the SAM to electronic communication. A defect, or
opening, is formed in the SAM allowing fluid to which the surface is exposed to
communicate electrically with the surface. Specifically, certain small sulfur containing
compounds having disruptive structures relative to the SAM as a whole were stably
incorporated into a SAM, and greater permeability to electrons was demonstrated. This

30   example demonstrates that a surface can be made electrically relatively conductive, and
then support cell growth.

A water-soluble ferrocene derivative was dissolved in the electrolyte solution: 100mM solution of ferrocenedicarboxylic acid in 500uM NaClO$_4$. The working electrode was a gold-coated electrode derivatized with a SAM comprised of varying amounts of 2-unit molecular wire (MFI). The height of the peak at a characteristic

5    ferrocene potential was plotted as a function of molecular wire density. As a negative control, a gold-coated electrode was derivatized with an insulating SAM comprised of 100% tri-ethylene glycol terminated thiol. This system was used to test the conductivity of electrodes modified with a panel of sulfur-containing compounds. The compounds were dissolved in DMF at 50% candidate compound and 50% tri-ethylene glycol

10   terminated thiol. Electrodes were derivatized as described in Example 1. SAMs were formed on gold chips from 500 micromolar triethylene glycol-terminated thiol and 500 micromolar of either mercaptobenzothiazole or 2-mercaptoethyl ether in DMF. The chips were clamped between a flat substrate and a 1 ml capacity silicon gasket. A solution of ferrocene dicarboxyllic acid was dissolved in 500 micro molar NaClO4 and

15   placed in the silicon gasket with a Ag/AgCl reference electrode and a Pt auxiliary electrode. The gold chip was connected as the working electrode. The system was analyzed by ACV. The magnitude of the current peaks, resulting from the ferrocene in solution communicating with the electrode, was an indicator of the ability of the trial compounds to make the SAM more permeable to electron flow by creating defects

20   within the SAM.

Example: Detection of Protein-protein interactions

This example demonstrates the utility of a colloid particle having an immobilized signaling entity and an immobilized protein. (See Fig. 18)

Histidine-tagged Glutathione-S-Transferase (GST-His) was attached to NTA-

25   SAM-coated colloids, displaying 40uM NTA-Ni and 100uM ferrocene-thiol. Commercially available magnetic beads presenting protein A were coated at 1/10 binding capacity with anti-GST antibody, added at a 1:5 ratio to the GST-colloids, and measured on a 50% MF-1 SAM-coated electrode, which was placed on top of a magnet. The magnet pulled the magnetic beads onto the electrode surface to form a

30   thick, visible precipitate. The GST-colloids were brought down to the electrode surface by the interaction with the GST-antibody on the magnetic beads to give a current peak at approximately 280mV. Two negative controls were run, one where GST was not

- 42 -

attached to the colloid surface, and another where the GST antibody was not attached to the magnetic beads. Neither negative control gave a current peak. Fig. 17 plots the results of this demonstration. Solid line represents interaction between GST-His-presenting colloids and anti-GST/Ab on magnetic beads. Open circles represent

5       magnetic beads presenting the antibody incubated with colloids that did not present GST. Closed circles represent beads not presenting the antibody, incubated with colloids that presented GST.

Prophetic Example: Massively Parallel Analysis of Protein-Protein Interactions: Elucidating the Interaction Map of the Human Proteome.

10       The following prophetic example describes how to perform massively parallel analysis of protein-protein interactions, which is particularly useful when proteins are as yet uncharacterized. Here, this method is used to elucidate the protein interaction map of the human proteome. A subset of, or the entire set of, proteins of the proteome is expressed with affinity tags to facilitate attachment of the expressed protein to sets of

15       particles. Particle-immobilized proteins are pooled together and allowed to interact. Interacting pairs are selected from the pooled mixture by a reiterative magnetic selection/dilution process. Following the selection/dilution process, the identity of interacting partners is determined. The selection step reduces the complexity of the problem by eliminating the need to analyze non-interacting proteins.

20       A diagram illustrating how two binding partners can be detected through magnetic recruitment is provided in Fig. 18. As shown, nanoparticle 9 is attached to electroactive signaling entity 10. An immobilized chemical or biochemical species 11 is also immobilized to the nanoparticle 9, as is oligo identifier 12 which is thereby associated with species 11. A second chemical or biochemical species 14 is

25       immobilized on magnetic bead 13. Second oligo identifier 15 is also immobilized with respect to magnetic bead 13 and thus is associated with second species 14. As binding occurs between species 11 and 14 (as shown in Fig. 18), a single hetero-particle forms that can then be magnetically drawn to sensing electrode 17 by electromagnet 16. Detection may be facilitated by signaling entity 10.

30       In a more specific example, proteins and their encoding DNA molecules are indirectly connected to each other by co-immobilizing both on a common particle or bead, wherein each particle (or bead) presents a single protein species and its encoding

- 43 -

DNA. Connecting the expressed protein to its encoding DNA expedites the identification of each set of interacting proteins after the selection/dilution process. Each protein and its encoding DNA are immobilized on two different kinds of particles: a recruitable particle and a signaling particle. The sizes of the particles are also

5    different such that smaller signaling particles can form satellites around each larger recruitable particle. In this example, each protein and its encoding DNA are immobilized on a single 4-10 micron magnetic bead as well as on a multitude of fluid suspendable nanoparticles, that are 4-40 nm in diameter and bear electroactive signaling entities. When a first species on a magnetic bead biologically interacts with a

10   second species on a signaling nanoparticle, the recruitable particle becomes "connected" to the signaling particle. These hetero-particle-complexes are then magnetically recruited to a sensing electrode (see Fig. 19), where they can deliver a signal, such as an electronic or electrochemical signal. To facilitate multiplexed analysis, the particles are magnetically attracted to an electrode array that has a number

15   of individually addressable electrode pads. Beneath each electrode pad is an individually controllable electromagnet, such that the magnetic field above each pad can be selectively turned off and on. However, magnetic particles that are not connected to signaling particles, which cannot deliver a signal, may also be recruited to the same electrode pad that delivers a positive signal. These non-signaling magnetic

20   beads are electromagnetically released from the pad and washed out of the interaction reservoir. Signaling nanoparticles that do not interact with species on magnetic particles remain in a homogeneous suspension and do not deliver a signal but are also washed out of the interaction reservoir. Both of these purging functions are accomplished by maintaining the magnetic fields beneath electrode pads that deliver a

25   positive signal, to retain interacting complexes, while the magnetic fields beneath pads that do not deliver a positive signal are driven to zero to release non-interacting magnetic beads. A port 150 (Fig. 19) is opened and fluid is washed out of the interaction reservoir carrying away non-interacting magnetic beads and non-interacting nanoparticles. The port is closed, all magnetic fields are driven to zero and more buffer

30   is added through a second port 160 along with mechanical agitation to resuspend and redistribute the particles. All the magnetic fields beneath all the electrode pads are turned on again and the selection/dilution process is repeated until, statistically, each

- 44 -

pad that delivers a positive signal contains a single magnetic bead bound to a multitude of signaling nanoparticles.

To determine the identity of interacting proteins, an array of magnetic pins, whose dimensions correspond to that of the electrode array is juxtaposed over the

5    electrode array and the magnetic fields beneath the entire electrode array are driven to zero such that each magnetic pin captures a single hetero-particle complex. The loaded pin array is dipped into a multi-well plate, of compatible dimensions, each well of which is filled with solution containing DNA amplification reagents. The respective encoding DNA sequences (immobilized on the nanoparticles and the bead) are

10    amplified by PCR or similar technique and sequenced to reveal the identity of each set of interacting proteins.

Prophetic Example: Preparation of Proteins for co-immobilization on surfaces that also present coding identifiers.

A DNA sequence that encodes each protein member of the proteome is inserted

15    into a bacterial protein expression vector. The expression vector carries an affinity tag, $(His)_6$ in this case, tandem repeats of Gal4 consensus sequences, and 2 sequences that flank the protein identification sequence, to which PCR primers can bind. The histidine-tag facilitates the attachment of the expressed protein to the particle. The Gal4 consensus sequences act to tether the encoding DNA to the particle via the

20    interaction between the recognition motif and a particle-immobilized yeast DNA binding domain, which in this case is a GST-Gal4 fusion protein. The DNA binding domain of Gal4 (aa' 1-100) binds to the consensus sequence CGGattAgAagcCgCCGAG and the GST binds to a glutathione moiety on the particle. Proteins are separately expressed in a cell-free translation system to reduce the

25    abundance of irrelevant proteins and cell debris. Following protein expression, each expression mixture contains the encoding DNA and the expressed protein. Each protein expression mixture is divided into 2 aliquots. A single magnetic bead (4-10 microns in diameter) is added to a first aliquot and a quantity of NTA-glutathione-SAM-coated nanoparticles is added to a second aliquot. Particles are pelleted and

30    washed to remove protein that is not particle-bound. Particles and beads are pooled together in subsets of 1000 species per pool and subjected to magnetic selection/dilution and electrochemical analysis. In this manner, unidentified proteins

can be bound to the same bead or particle that also binds the corresponding encoding DNA.

In examples were colloid particles (nanoparticles) are used, they can be prepared as described in the above-referenced international patent publication nos. WO
00/43791 and WO 00/34783

Prophetic example: Electrochemical Analysis

An electrochemical analyzer from CH Instruments (Austin, TX) is used to detect interactions between species immobilized on magnetic beads and species immobilized on colloids that also bear redox active metals. The instrument is modified to facilitate multiplexed detection. In this case, the redox active metals are ferrocene derivatives. Pads of the electrode array are individually addressable and act as the working electrode. In this case, the pads are gold-coated and derivatized with conductive self-assembled monolayers. A Ag vs. Ag/Cl reference electrode is used with a Pt auxiliary. Electrodes are scanned using Alternating Current Voltammetry (ACV) with a 25mV overpotential at a frequency of 10 Hz..

Prophetic Example: Design of Electrode Array Sandwiched between Individually addressable Electromagnets

Electrode arrays 100 (Fig. 19) having 300-500 electrode pads 110 are constructed by plating gold over $Ni^{++}$. Electrode pads are 50-500 microns on edge and are sandwiched between sets of individually addressable Helmholtz electromagnets 120 such that a magnetic field gradient can be generated to recruit, then hold magnetic beads at the pad surface (see Fig. 19). The direction of the current is reversed to drive the magnetic field to zero when it is desirable to release magnetic beads from the surface to wash away or redistribute. To ensure that the interaction reservoir is thermally isolated from the electromagnet arrays so that heat does not denature proteins in solution, a layer of insulative material 130 is placed between the electromagnets and the interaction reservoir 140.

Prophetic Example: Calculation of Protein Sets and Electrode Pad Number

To generate the protein interaction map of the entire proteome, one needs to divide the proteome into subsets, which are then tested for interaction with every other subset. Assuming that there are about 50,000 proteins of interest, the proteome is divided into 50 sets of 1000 proteins each. Each group of 1000 proteins is then tested for interaction

- 46 -

with every other group of 1000, resulting in 50X50 matrix or 2500 separate experiments. The number of proteins in each subset determines the number of electrode pads in each array. If we assume that each protein has a single binding partner, then each protein has a 1/50 chance of finding that partner when tested for interaction with one of the 50 subsets of proteins. However, each protein probably has on average 5 relevant binding partners, increasing the probability of finding a binding partner within a subset to 1/10. That means that for 2000 proteins in one pooled interaction mix, 200 will deliver a positive signal, which implies that the electrode array should have 300-500 pads. Low-level signals from non-specific binding events are minimal because of competitive inhibition by relevant binders. However, the occurrence of false positives is minimized when a signal threshold is set, wherein signals below the threshold are counted as negatives. Relative affinities are determined by comparison of the degree of interaction of a first binding species and a second binding species with a third target protein to which the first and second bind.

Prophetic Example:  Determining the Binding Partners of a Single Target Protein with a Large Pool of Candidate Binding Partners

This prophetic example describes how to identify proteins, from a large pool of putative binding partners, which interact with a single target protein. Proteins from the large pool are prepared as described above in Example 30, and immobilized only on signaling nanoparticles. The target protein is immobilized on a set of magnetic beads. Because the identity of the target protein is known, it is not necessary to co-immobilize its encoding DNA. Interacting partners are selected by electrochemical analysis as described above.

Prophetic Example:  Selection of Interacting Protein Partners by FACS Analysis

This example describes how to identify the binding partners of a single target protein. The target protein is immobilized on a set of beads that are 4-25 microns in diameter. Putative binding partners, which may be prepared as described above or generated from a cDNA library, are co-immobilized along with their encoding DNA onto nanoparticles that bear fluorescent signaling moieties. When a bead-immobilized protein interacts with a species immobilized on nanoparticles, the bead becomes decorated with fluorescent nanoparticles and can be isolated by FACS (fluorescent

- 47 -

activated cell sorting) analysis after which the attached DNA of each interacting species is sequenced to identify the binding partners.

Those skilled in the art would readily appreciate that all parameters listed herein are meant to be exemplary and that actual parameters will depend upon the specific application for which the methods and apparatus of the present invention are used. It is, therefore, to be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, the invention may be practiced otherwise than as specifically described.

What is claimed is:

1.    A method comprising:

allowing a chemical or biological species, immobilized relative to a surface, to participate in a chemical or biological interaction; and

determining participation of the chemical or biological species in the chemical or biological interaction by identifying an oligonucleotide identifier associated with the surface.

2.    A method in claim 1, wherein the surface comprises gold.

3.    A method as in claim 2, wherein the surface is a surface of a gold colloid particle.

4.    A method as in claim 3, wherein the chemical or biological species is immobilized relative to the surface via a self-assembled monolayer.

5.    A method as in claim 1, wherein the chemical or biological species is fastened to the surface via a metal binding tag/metal/chelate linkage.

6.    A method as in claim 1, wherein, during the allowing step, the oligonucleotide identifier is fastened to the surface, the determining step comprising separating the oligonucleotide identifier from the surface and then identifying the oligonucleotide identifier.

7.    A method as in claim 6, wherein, during the allowing step, the oligonucleotide identifier is fastened to the surface via a self-assembled monolayer.

8.    A method as in claim 6, comprising identifying the oligonucleotide identifier via fluorescent sequencing.

9.    A method as in claim 1, the allowing step comprising allowing a first species, fastened to a first surface, to biologically bind to a second species fastened to a second surface;

- 49 -

determining immobilization of the first surface relative to the second surface; and

identifying the species fastened to the second surface by identifying an oligonucleotide identifier which was fastened to the surface of the second article during the allowing step.

10.    A method as in claim 9, wherein each of the first and second articles is a colloid particle.

11.    A method as in claim 1, comprising identifying the oligonucleotide identifier by identifying a complementary oligonucleotide having a first portion complementary to the oligonucleotide identifier and a second portion complementary to a second oligonucleotide identifier.

12.    A method as in claim 1, comprising allowing a first chemical or biological species, immobilized relative to a surface of a first article, to chemically or biologically interact with a second chemical or biological species, immobilized relative to a surface of a second article; and

determining the chemical or biological interaction by identifying an interaction hybridization identifier that is complementary to a combination of a first oligonucleotide identifier fastened to the surface of the first article and a second oligonucleotide identifier fastened to the surface of the second article.

13.    A method as in claim 12, comprising providing a first colloid particle, a first species fastened to the first colloid particle, and a first oligonucleotide identifier fastened to the first colloid particle, a second colloid particle, a second species fastened to the second colloid particle, and a second oligonucleotide identifier fastened to the second colloid particle;

allowing the first and second species to biologically bind, thereby immobilizing the first and second colloid particles relative to each other and bringing the first oligonucleotide identifier into proximity with the second oligonucleotide identifier;

- 50 -

exposing the first and second oligonucleotide identifiers to an interaction

hybridization identifier that is complementary to the combination of the first and

second oligonucleotide identifiers and allowing the interaction hybridization identifier

to bind to the first and second oligonucleotide identifier; and

5         identifying the interaction hybridization identifier thereby identifying the first

and second oligonucleotide identifiers and thereby identifying the biological binding.


14.     A method as in claim 13 comprising, prior to the identifying step, de-activating

any non-hybridized oligonucleotide.

10

15.     A kit comprising:

an article having a surface;

a chemical or biological species, able to participate in a chemical or biological

interaction, fastened to or adapted to be fastened to the surface; and

15         an oligonucleotide identifier fastened to or adapted to be fastened to the surface.


16.     A kit as in claim 15, wherein the article is a colloid particle.


17.     A kit as in claim 15, wherein the article is a first article, the chemical or

20   biological species is a first chemical or biological species, and the oligonucleotide

identifier is a first oligonucleotide identifier, further comprising:

a second article having a surface;

a second chemical or biological species, able to participate in a chemical or

biological interaction, fastened to or adapted to be fastened to the second surface; and

25         a second oligonucleotide identifier fastened to or adapted to be fastened to the

second surface.


18.     A kit as in claim 17, wherein each of the first and second articles is a colloid

particle.

30

19.     A kit as in claim 17, wherein each of the first and second chemical or biological species is fastened to or adapted to be fastened to the first or second surface, respectively, via a metal binding tag/metal/chelate linkage.

5    20.     A kit as in claim 19, wherein each of the first and second chemical or biological species and first and second oligonucleotide identifiers is fastened to or adapted to be fastened to the first or second surface via a self-assembled monolayer-forming species.

21.     A kit as in claim 17, further comprising an interaction hybridization identifier
10   that is complementary to a combination of the first and second oligonucleotide identifier.

22.     A kit comprising a plurality of particles each carrying a chemical or biological functionality allowing it to fasten to a binding partner, and each carrying an identical
15   oligonucleotide linker constructed for attachment to a complementary oligonucleotide fastened to an oligonucleotide identifier.

23.     A composition comprising:
        a chemical or biological species, able to participate in a chemical or biological
20   interaction;
        a linker species that is not a ribosome; and
        an oligonucleotide identifier, wherein each of the chemical or biological species and the oligonucleotide identifier is fastened to or adapted to be fastened to the linker species.
25

24.     A composition comprising:
        a protein;
        an linker species that is not a ribosome; and
        an oligonucleotide identifier that encodes for the protein, wherein each of the
30   protein and the oligonucleotide identifier is immobilized relative to, or adapted to be immobilized relative to, the linker species.

- 52 -

25.    A composition as in claim 24, wherein the linker species is a nanoparticle and each of the protein and the oligonucleotide identifier is immobilized relative to or adapted to be immobilized relative to a surface of the nanoparticle.

26.    A composition as in claim 24, wherein the linker species is a chip and each of the protein and the oligonucleotide identifier is immobilized relative to or adapted to be immobilized relative to a surface of the chip.

27.    A composition as in claim 24, wherein the linker species is a polymer and each of the protein and the oligonucleotide identifier is immobilized relative to or adapted to be immobilized relative to the polymer.

28.    A composition as in claim 24, wherein the linker species is a dendrimer and each of the protein and the oligonucleotide identifier is immobilized relative to or adapted to be immobilized relative to the dendrimer.

29.    A composition as in claim 24, wherein the linker species is a RNA binding protein and each of the protein and the oligonucleotide identifier is immobilized relative to or adapted to be immobilized relative to the RNA binding protein.

30.    A composition as in claim 24, wherein the linker species is a DNA binding protein and each of the protein and the oligonucleotide identifier is immobilized relative to or adapted to be immobilized relative to the DNA binding protein.

31.    A composition as in claim 24 further comprising a chimeric oligo solution that is complimentary to both the oligonucleotide identifier that encodes for the protein and a second oligonucleotide identifier that encodes for a binding partner of the protein.

32.    A kit comprising:
         a surface;
         a protein immobilized relative to the surface or adapted to be immobilized relative to the surface; and

- 53 -

an oligonucleotide identifier that codes for the protein, immobilized relative to the surface or adapted to be immobilized relative to the surface.

33.     A kit as in claim 32, wherein at least a portion of the surface is coated with a self-assembled monolayer.

34.     A kit as in claim 33, wherein each of the oligonucleotide identifier and protein is immobilized or adapted to be immobilized relative to the common surface via the self-assembled monolayer.

35.     A kit as in claim 32, wherein the surface is a surface of a recruitable particle.

36.     A kit as in claim 32, wherein the surface is a surface of a magnetic bead.

37.     A kit as in claim 32, wherein the surface is a surface of a colloid particle.

38.     A kit as in claim 32, wherein the surface is a surface of a chip.

39.     A kit as in claim 32, wherein the oligonucleotide identifier is hybridized or hybridizable to an oligonucleotide sequence that is fastened to or adapted to be fastened to the surface.

40.     A kit as in claim 39, wherein the oligonucleotide identifier is hybridized or hybridizable to an oligonucleotide sequence that forms part of a self-assembled monolayer fastened to the surface.

41.     A kit as in claim 32, wherein the oligonucleotide identifier comprises plasmid DNA.

42.     A kit as in claim 32, wherein the oligonucleotide identifier comprises a protein expression vector.

- 54 -

43.     A kit as in claim 32, wherein the oligonucleotide identifier comprises linear DNA.

44.     A kit as in claim 43, wherein the oligonucleotide identifier comprises a protein expression template.

45.     A kit as in claim 43, wherein the oligonucleotide identifier comprises a product of a polymerase chain reaction.

46.     A kit as in claim 45, wherein the oligonucleotide identifier comprises a protein expression template.

47.     A kit as in claim 32, wherein the oligonucleotide identifier comprises a protein expression template.

48.     A kit as in claim 32, wherein the oligonucleotide identifier is immobilized to or adapted to be immobilized to the surface via a nucleic acid binding protein.

49.     A kit as in claim 32, wherein the oligonucleotide identifier is modified to facilitate attachment to the surface via a recognition protein.

50.     A kit as in claim 32, wherein the oligonucleotide identifier is biotinylated to facilitate attachment to the surface via streptavidin.

51.     A kit as in claim 47, wherein the oligonucleotide identifier is immobilized to or adapted to be immobilized to the surface via a DNA binding protein.

52.     A kit as in claim 32, further comprising a signaling entity immobilized relative to or adapted to be immobilized relative to at least one of the oligonucleotide identifier and protein.

- 55 -

53.    A kit as in claim 52, wherein the oligonucleotide identifier is modified to include a signaling entity.

54.    A kit as in claim 53, wherein the oligonucleotide identifier is generated via PCT using primers modified with signaling entities.

55.    A kit as in claim 32, wherein the oligonucleotide identifier comprises PCR sites.

56.    A kit as in claim 32, wherein the protein immobilized relative to the surface or adapted to the immobilized relative to the surface is expressed off of the oligonucleotide identifier.

57.    A kit comprising:
       a polymer or dendrimer;
       a protein immobilized relative to the polymer or dendrimer or adapted to be immobilized relative to the polymer or dendrimer; and
       an oligonucleotide identifier that codes for the protein, immobilized relative to the polymer or dendrimer or adapted to be immobilized relative to the polymer or dendrimer.

58.    A composition comprising:
       a protein and an oligonucleotide identifier that codes for the protein, immobilized relative to each other or adapted to be immobilized relative to each other.

59.    A kit comprising:
       a protein and an oligonucleotide identifier that codes for the protein, immobilized relative to each other or adapted to be immobilized relative to each other; and
       an entity carrying immobilized thereto a binding partner of the protein.

60.    A kit as in claim 59, wherein the entity is capable of carrying immobilized thereto a plurality of binding partners of the protein.

61.     A kit as in claim 59, wherein the entity carries immobilized thereto a plurality of binding partners of the protein.

62.     A kit as in claim 59, wherein the entity is a recruitable particle.

63.     A kit as in claim 59, wherein the entity is a magnetic bead.

64.     A kit as in claim 59, wherein the entity is a colloid particle.

65.     A kit as in claim 59, wherein the entity is a surface of a chip.

66.     A kit as in claim 59, further comprising an oligonucleotide identifier immobilized to or adapted to be immobilized to the binding partner.

67.     A kit as in claim 59, wherein the protein is a fusion protein.

68.     A kit as in claim 66, wherein the protein comprises a binding partner and an affinity tag.

69.     A method comprising:
        expressing a protein with an oligonucleotide;
        immobilizing the protein and the oligonucleotide relative to each other.

70.     A method as in claim 69, wherein each of the oligonucleotide identifier and protein is immobilized or adapted to be immobilized relative to a common surface.

71.     A method as in claim 70, wherein at least a portion of the surface is coated with a self-assembled monolayer.

72.     A method as in claim 71, wherein each of the oligonucleotide identifier and protein is immobilized or adapted to be immobilized relative to the common surface via the self-assembled monolayer.

5     73.     A method as in claim 69, wherein the surface is a surface of a recruitable particle.

74.     A method as in claim 73, wherein the surface is a surface of a magnetic bead.

10     75.     A method as in claim 73, wherein the surface is a surface of a colloid particle.

76.     A method as in claim 69, wherein the surface is a surface of a chip.

77.     A method as in claim 69, wherein each of the oligonucleotide identifier and
15     protein is immobilized or adapted to be immobilized relative to a common polymer.

78.     A method as in claim 69, wherein each of the oligonucleotide identifier and protein is immobilized or adapted to be immobilized relative to a common dendrimer.

20     79.     A method as in claim 70, wherein the oligonucleotide identifier is hybridized or hybridizable to an oligonucleotide sequence that is fastened to or adapted to be fastened to the surface.

80.     A method as in claim 69, wherein the oligonucleotide identifier comprises
25     plasmid DNA.

81.     A kit as in claim 69, wherein the oligonucleotide identifier comprises a protein expression vector.

30     82.     A method as in claim 69, wherein the oligonucleotide identifier comprises linear DNA.

- 58 -

83.     A kit as in claim 82, wherein the oligonucleotide identifier comprises a protein expression template.


84.     A kit as in claim 82, wherein the oligonucleotide identifier comprises a product 5    of a polymerase chain reaction.


85.     A kit as in claim 84, wherein the oligonucleotide identifier comprises a protein expression template.


10    86.     A method as in claim 70, wherein the oligonucleotide identifier is immobilized to or adapted to be immobilized to the surface via a nucleic acid binding protein.


87.     A method as in claim 86, wherein the oligonucleotide identifier is immobilized to or adapted to be immobilized to the surface via a DNA binding protein.

15

88.     A method as in claim 69, wherein the oligonucleotide identifier and protein are immobilized to or adapted to be immobilized relative to each other in the absence of a common surface to which each is immobilized.


20    89.     A method as in claim 88, further comprising a signaling entity immobilized relative to or adapted to be immobilized relative to at least one of the oligonucleotide identifier and protein.


90.     A method as in claim 88, further comprising exposing the protein to an entity 25    carrying immobilized thereto a binding partner of the protein.


91.     A method as in claim 90, comprising exposing the protein to an entity capable of carrying immobilized thereto a plurality of binding partners of the protein.


30    92.     A method as in claim 90, comprising exposing the protein to an entity carrying immobilized thereto a plurality of binding partners of the protein.

- 59 -

93.     A method as in claim 90, wherein the entity is a recruitable particle.

94.     A method as in claim 90, wherein the entity is a magnetic bead.

5    95.     A method as in claim 90, wherein the entity is a colloid particle.

96.     A method as in claim 90, wherein the entity is a surface of a chip. .

97.     A method as in claim 90, further comprising an oligonucleotide identifier
10    immobilized to or adapted to be immobilized to the binding partner.

98.     A method as in claim 69, wherein the protein is a fusion protein.

99.     A method as in claim 98, wherein the protein comprises a binding partner and
15    an affinity tag.

100.    A method as in claim 98, further comprising a signaling entity immobilized
relative to the protein and the oligonucleotide identifier.

20    101.    A method as in claim 100, wherein the signaling entity is part of the fusion
protein.

102.    A method comprising:
        allowing a chemical or biological species to participate in a chemical or
25    biological interaction; and
        determining participation of the chemical or biological species in the chemical
or biological interaction by identifying an oligonucleotide identifier, wherein the
oligonucleotide identifier encodes the chemical or biological species.

30    103.    Generating a library of nucleic acids that contain components of a cDNA library
and a functionality to facilitate binding to a surface.

104. Generating a library of nucleic acids that contain components of a cDNA library and a functionality the products of which are used in an in vitro assay.

105. Generating a library of nucleic acids that contain components of a cDNA library and sequences to which nucleic acid binding proteins bind.

106. Generating a library of plasmids that contain components of a cDNA library and a functionality to facilitate binding to a surface.

107. Generating a library of plasmids that contain components of a cDNA library and a functionality the products of which are used in an in vitro assay.

108 Generating a library of plasmids that contain components of a cDNA library and sequences to which nucleic acid binding proteins bind.

109. Generating a library of nucleic acids or plasmids that contain components of a cDNA library, sequences that encode a DNA binding domain and sequences to which the encoded DNA binding domain binds.

110. Generating a library of nucleic acids or plasmids that contain components of a cDNA library, sequences that encode a DNA binding domain and sequences to which the encoded DNA binding domain binds, wherein the binding motif sequences are not in proximity to a reporter gene.

111. A kit comprising:

at least one colloid particle;

at least one magnetic bead;

at least one protein recognition motif adapted for immobilization to the at least one colloid particle; and

an uncharacterized protein or drug adapted for immobilization to the at least one bead.

- 61 -

112. The kit of claim 111 further comprising DNA adapted for immobilization to the at least one bead.

113. The kit of claim 112 wherein the DNA encodes for the uncharacterized protein.

5

114. A method comprising:

exposing a plurality of colloid particles, each carrying an immobilized protein recognition motif, to a bead carrying an immobilized, uncharacterized protein or drug; and

10        determining immobilization of at least one particle to the bead via interaction between the protein recognition motif and the uncharacterized protein or drug.

115. A method as in claim 114, further comprising determining the identity of the uncharacterized protein or drug by determining which protein recognition motifs it

15    binds to.

116. The method of claim 115 wherein the presence of an unknown protein or drug is determined by detecting an identifier attached to the bead, the identifier corresponding to the uncharacterized protein or drug.

20

117.        The method of claim 116 wherein the identifier is DNA.

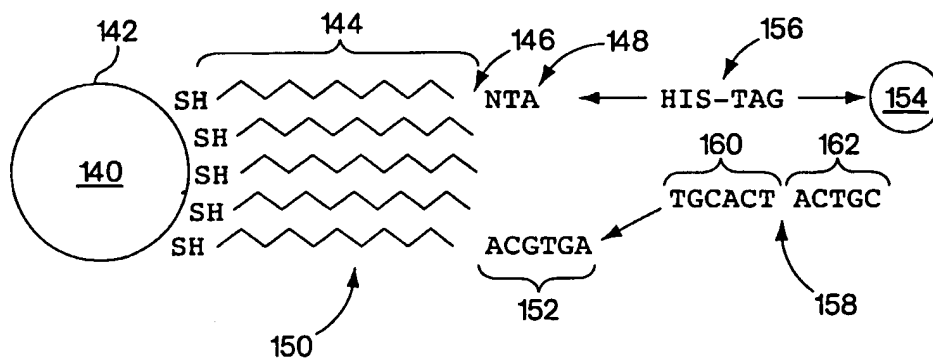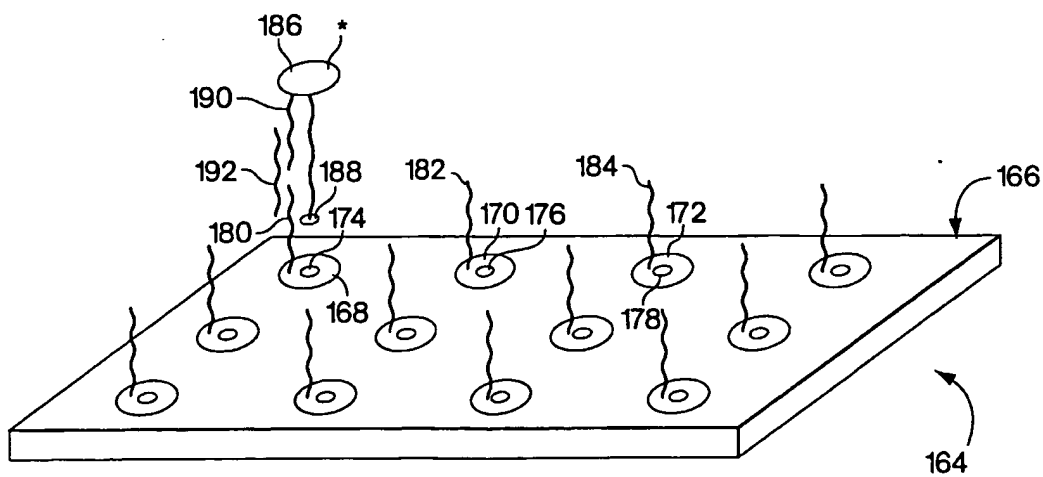118.        The method of claim 117 wherein the DNA encodes for the uncharacterized protein.
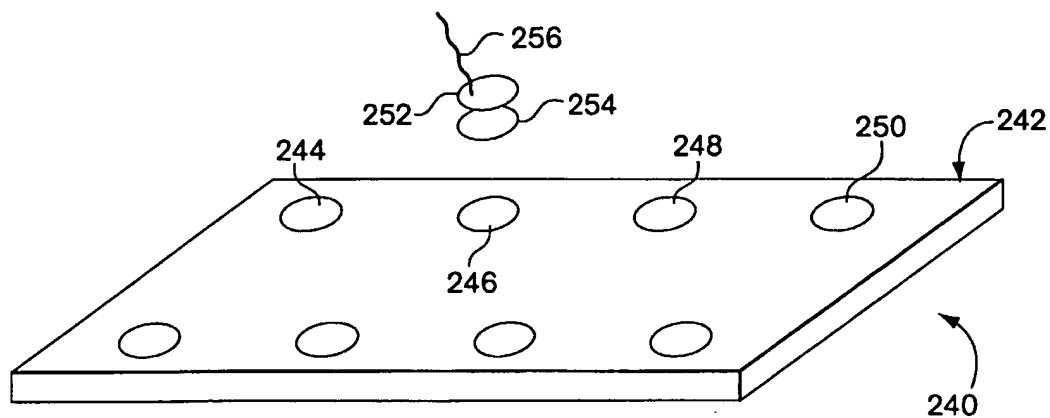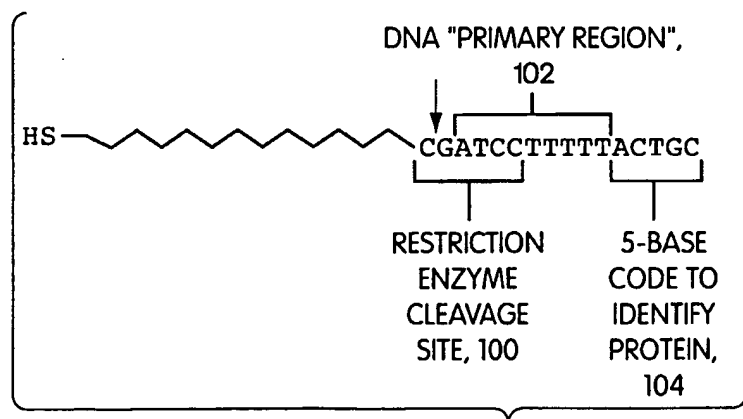
25

Fig. 1



Fig. 2

**Fig. 3**



**Fig. 4**

FOLD COMPLEMENTARY DNA TO "DNA PRIMING REGION" +
SEQUENCE USING STANDARD PCR METHODS:

COMPLEMENTARY OLIGO, 122 ⌐

                              ⌐CTAGGAAAAA¬  SEQUENCE
                               GATCCTTTTTACTGC        →
                               ‿‿‿‿‿‿ ‿‿‿‿
                                 102      104

MATCH UP RESULTING SEQUENCE DATA WITH
RECORDS KEPT THAT CONNECT PROTEIN
IDENTIFY TO SEQUENCE:

        ACTGC = PROTEIN # 120
        ‿‿‿       (SPECIES)
         104

## Fig. 5

110

SAM, 108

SH ∿∿∿∿∿∿∿∿
SH ∿∿∿∿∿∿∿∿              100              104
SH ∿∿∿∿∿∿∿∿           ⌐            ¬    ⌐     ¬
COLLOID,  SH ∿∿∿∿∿∿∿∿   CGATCCTTTTTACTGC
106       SH ∿∿∿∿∿∿∿∿            ‿‿‿‿
SH ∿∿∿∿∿∿∿∿               102
SH ∿∿∿∿∿∿∿∿        ⌐COOH
                   COOH-Ni-HIS-TAG—  PROTEIN,
SH                 ⌐COOH              112
           ↙      ‿‿‿‿‿‿‿‿‿
                   NTA    114

OLIGO-THIOL IS INCORPORATED INTO SAM
ALONG WITH NTA ON COLLOID. HIS-TAGGED
PROTEIN IS ATTACHED TO COLLOID VIA NTA-Ni

## Fig. 6

PROTEIN #1

COLLOID,
106

112  120

COLLOID,
118

PROTEIN #2

PROTEIN, BOUND TO
COLLOID IS ALLOWED TO
PARTICIPATE IN BINDING
ASSAYS.
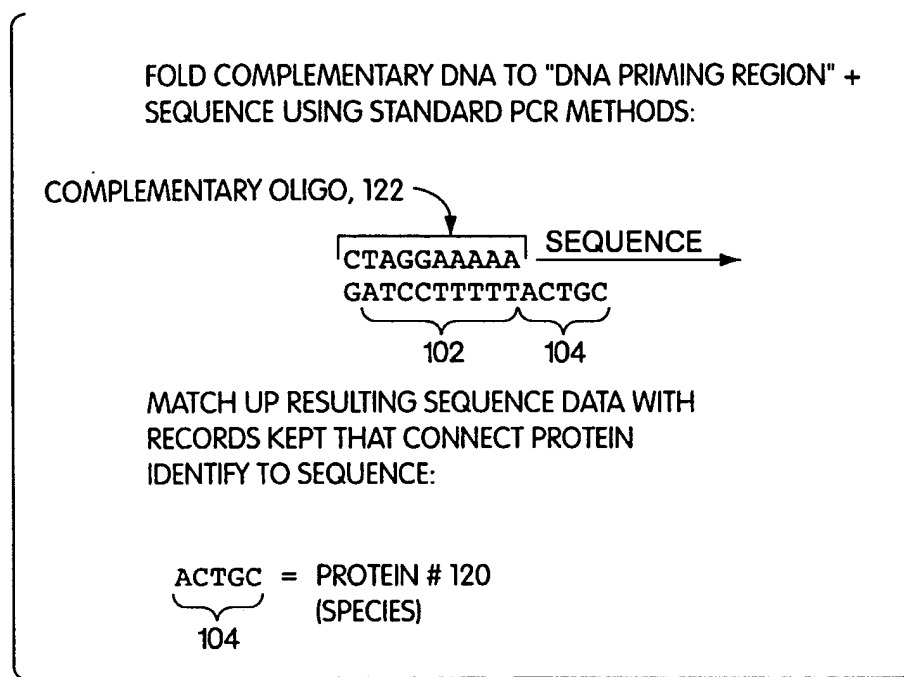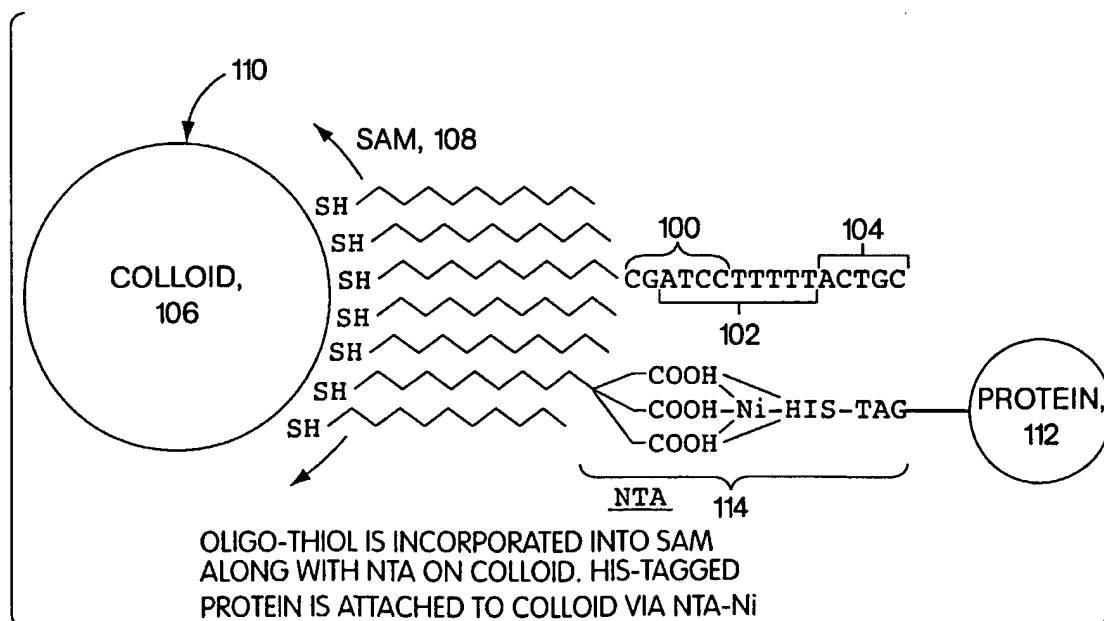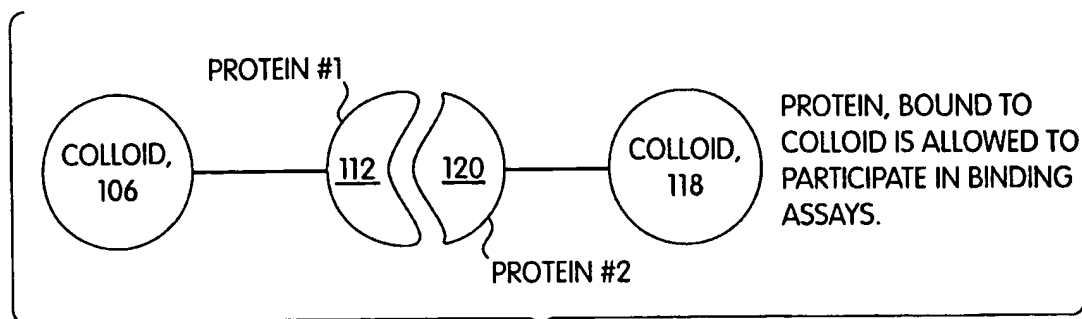
## Fig. 7

TO UNCOVER THE IDENTITY OF PROTEIN AFTER
THE ASSAYS ARE COMPLETED, CLEAVE THE
DNA PORTION OF THE DNA-THIOL BY ADDITION
OF A RESTRICTION ENZYME:

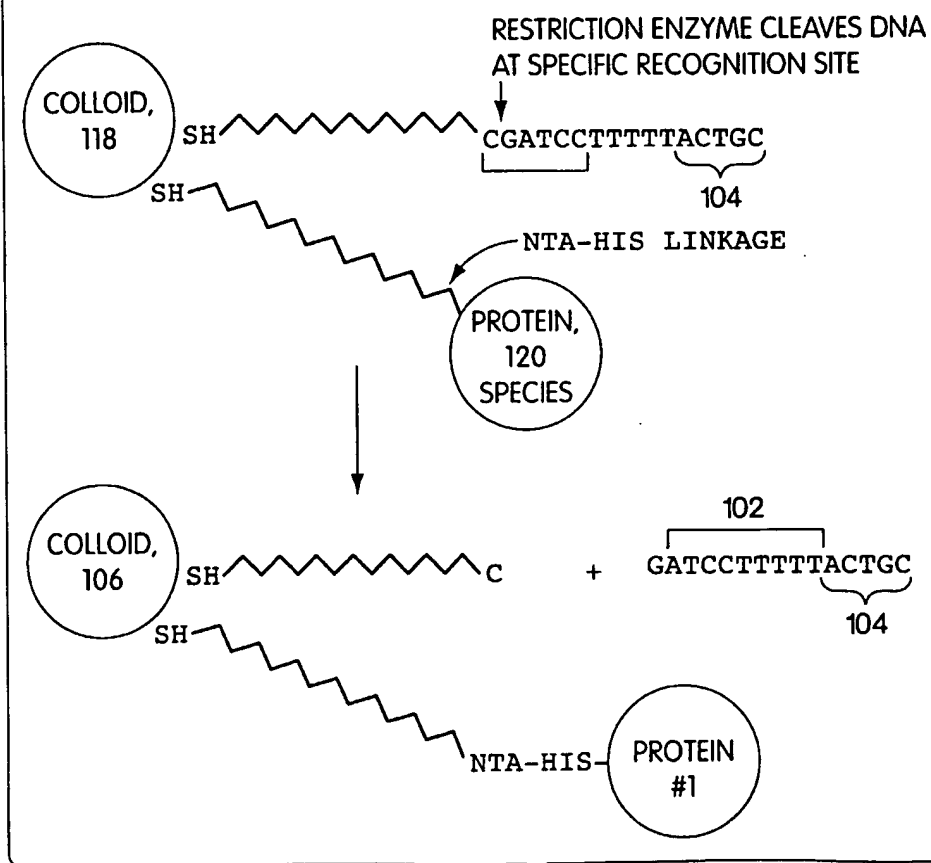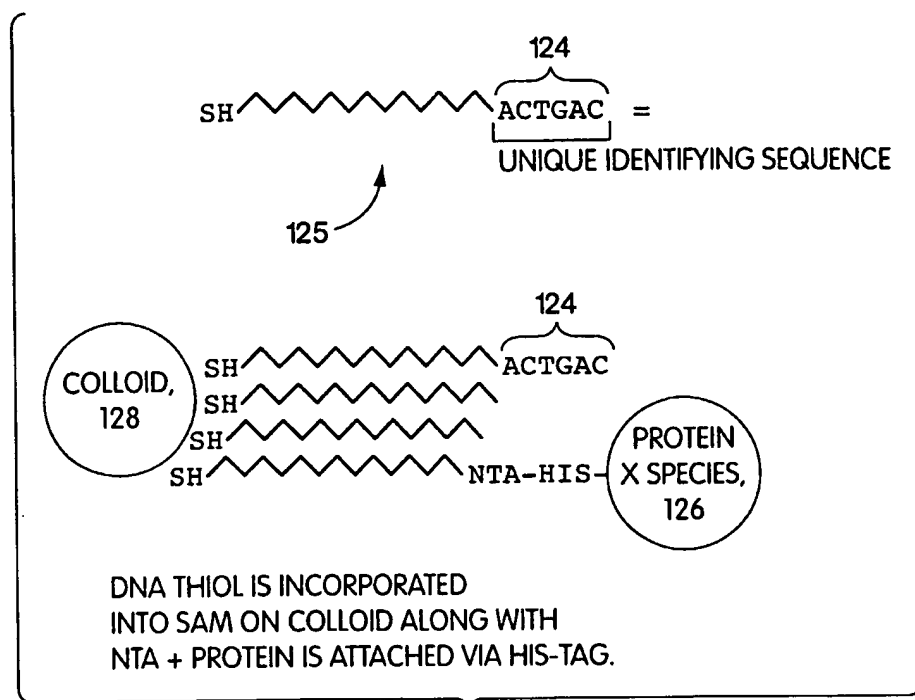RESTRICTION ENZYME CLEAVES DNA
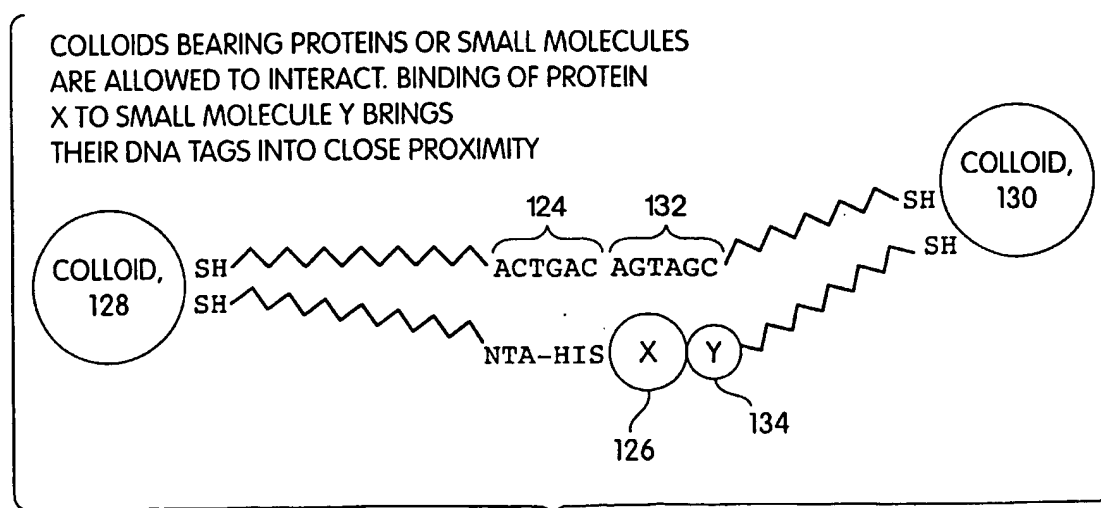AT SPECIFIC RECOGNITION SITE

COLLOID,
118

SH

CGATCCTTTTTACTGC

104

SH

NTA-HIS LINKAGE

PROTEIN,
120
SPECIES

COLLOID,
106

SH

C

+

102

GATCCTTTTTACTGC

104

SH

NTA-HIS

PROTEIN
#1

## Fig. 8

5/11

124
SH⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋‿ACTGAC‿ =
                        UNIQUE IDENTIFYING SEQUENCE

125

124

COLLOID,   SH⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋‿ACTGAC
128        SH⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋
           SH⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋                    PROTEIN
           SH⟍⟋⟍⟋⟍⟋⟍⟋⟍⟋‿NTA-HIS‿          X SPECIES,
                                                126

DNA THIOL IS INCORPORATED
INTO SAM ON COLLOID ALONG WITH
NTA + PROTEIN IS ATTACHED VIA HIS-TAG.

## Fig. 9

COLLOIDS BEARING PROTEINS OR SMALL MOLECULES
ARE ALLOWED TO INTERACT. BINDING OF PROTEIN
X TO SMALL MOLECULE Y BRINGS
THEIR DNA TAGS INTO CLOSE PROXIMITY

124    132                      COLLOID,
                                 130
COLLOID,   SH⟍⟋⟍⟋⟍⟋⟍⟋‿ACTGAC AGTAGC⟍⟋⟍⟋⟍⟋‿SH
128        SH⟍⟋⟍⟋⟍⟋⟍⟋                    ⟍⟋⟍⟋⟍⟋‿SH
                        NTA-HIS( X )( Y )
                                126    134

## Fig. 10

COMPLEMENTARY SEQUENCES TO OLIGO IDENTIFIER DNA
ARE ADDED + ALLOWED TO BIND.

CACCGTATTAGT

138

REMOVE BY
DNAase

140   REMOVE BY
       DNAase

GTACGCCGTTGT

136

TGACTGTCATCG
ACTGACAGTAGC

124      132

COLLOID
X,
128

COLLOID
Y,
130

SH

SH

X    Y

126        134

## Fig. 11

COLLOID
Z

GAATCA

REMOVE BY
DNAase

Z

SINGLE-STRANDED DNAase IS ADDED TO REMOVE (OR
"CHEW UP") ANY NON-HYBRIDIZED DNA.

## Fig. 12

RESULT

136

TGACTGTCATCG
ACTGACAGTAGC

124      132

COLLOID
Y,
130

COLLOID
X,
128

COLLOID
Z

X    Y

126        134

Z

## Fig. 13

COMPLEMENTARY DNA IS DENATURED AND SEQUENCED.

136

-TGACTGTCATCG-

ACTGACAGTAGC

COLLOID
X,
128

124      132

X      X

126      134

COLLOID
Y,
130

## Fig. 14

RESULTING SEQUENCE CONTAINS THE
UNIQUE DNA CODES OF THE TWO BINDING
PARTNERS, X + Y:

136

ACTGACAGTAGC

=      ACTGAC      +      AGTAGC

124                132

UNIQUE
SEQUENCE
FOR
PROTEIN X
(SPECIES, 126)

UNIQUE
SEQUENCE
FOR
PROTEIN Y
(SPECIES, 134)

PROTEIN X + PROTEIN Y MUST BE BINDING PARTNERS.

## Fig. 15

Tech:ACV
File:negconbb.bin

init E (V) = 0
Final E (V) = 0.8
Incr E (V) = 0.008
Amplitude (V) = 0.025
Frequency (Hz) = 10
Sample Period (s) = 1
Quiet Time (s) = 2
Sensitivity (A/V) = 1e-5

◇   negconbb.bin

—   posconb.bin

Fig. 16

Tech:ACV
File:sb062_007bb

init E (V) = 0.1
Final E (V) = 0.7
Incr E (V) = 0.008
Amplitude (V) = 0.025
Frequency (Hz) = 10
Sample Period (s) = 1
Quiet Time (s) = 2
Sensitivity (A/V) = 5e-4

— sb062_007bb

o   sb062_012bb.bin

Fig. 17

## Fig. 18

Fig. 19